

MILITARY OPERATIONS RESEARCH SOCIETY



MORS Workshop

Improving Defense Analysis Through Better Data Practices

25-27 March 2003

Institute for Defense Analyses

Alexandria, Virginia

Co-Chairs:

Thomas L. Allen and James N. Bexfield, FS

UNCLASSIFIED
Approved for Public Release

1703 N. Beauregard Street ♦ Suite 450 ♦ Alexandria Virginia 22311-1717

(703) 933-9070 ♦ FAX: (703) 933-9066 ♦ email: morsoffice@mors.org

URL: <http://www.mors.org>

20050830 004

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person should be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)

3 December 2004

2. REPORT TYPE

Workshop Report

3. DATES COVERED (From – To)

25-27 March 2003

4. TITLE AND SUBTITLE

**MORS Workshop Report
Improving Defense Analysis Through Better Data Practices**

5a. CONTRACT NUMBER

N00014-99-C-0113

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

**Thomas L. Allen and James N. Bexfield, FS
Corrina Ross-Witkowski, Editor/Publisher**

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

**Military Operations Research Society, 1703 N. Beauregard St,
Suite 450, Alexandria, VA 22311**

8. PERFORMING ORGANIZATION REPORT
NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

**Distribution Statement A:
Approved for Public Release; Distribution Unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

**This publication is the report of the Workshop on Improving Defense Analysis Through Better
Data Practices.**

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION:
UNCLASSIFIED

a. REPORT

Unclassified

b. ABSTRACT

Unclassified

c. THIS PAGE

Unclassified

17. LIMITATION
OF ABSTRACT

Unlimited

18. NUMBER OF
PAGES

**1- 186and
Appendixes
A-1-A-30
B-1-B-4
C-1-C-10**

19a. NAME OF RESPONSIBLE PERSON

Corrina Ross-Witkowski

19b. TELEPHONE NUMBER (include area
code)

**703-933-9070
FAX: 703-933-9066**

MILITARY OPERATIONS RESEARCH SOCIETY



MORS Workshop

Improving Defense Analysis Through Better Data Practices

25-27 March 2003

**Institute for Defense Analyses
Alexandria, Virginia**

Co-Chairs:

Thomas L. Allen and James N. Bexfield, FS

UNCLASSIFIED

Approved for Public Release

**1703 N. Beauregard Street ♦ Suite 450 ♦ Alexandria Virginia 22311-1717
(703) 933-9070 ♦ FAX: (703) 933-9066 ♦ email: morsoffice@mors.org
URL: <http://www.mors.org>**

The Military Operations Research Society (MORS)

The purpose of the Military Operations Research Society (MORS) is to enhance the quality and effectiveness of classified and unclassified military operations research. To accomplish this purpose, the Society provides media for professional exchange and peer criticism among students, theoreticians, practitioners and users of military operations research. These media consist primarily of the traditional annual MORS Symposia (classified), their published abstracts or proceedings, special mini-symposia, workshops, colloquia and special purpose monographs and other publications. MORS publishes two quarterly periodicals, *PHALANX* and *Military Operations Research*. *PHALANX* is the MORS bulletin and *Military Operations Research* is a refereed journal. The forum provided by these media is intended to display the state of the art, to encourage consistent professional quality, to stimulate communication and interaction between practitioners and users, and to foster the interest and development of students of operations research. The Military Operations Research Society neither makes nor advocates official policy, nor does it attempt to influence the formulation of policy. Matters discussed or statements made during the course of its meetings or printed in its publications represent the positions of the individual participants and authors and not of the Society.

The Military Operations Research Society is operated by a Board of Directors consisting of 30 members, 28 of whom are elected by vote of the Board to serve a term of four years. The persons nominated for the board generally are individuals who have attained recognition and prominence in the field of military operations research, and who have demonstrated an active interest in the programs and activities of MORS. The remaining two members of the Board of Directors are the Immediate Past President who serves by right and the Executive Vice President who serves as a consequence of his position. A limited number of Advisory Directors are appointed from time to time, usually for a one-year term, to perform some particular function.

MORS is sponsored by:

- The Deputy Under Secretary of the Army (Operations Research)
- The Director, Assessment Division, Office of the Chief of Naval Operations
- Director, Air Force Studies and Analyses Agency
- The Commanding General, Marine Corps Combat Development Command
- The Director of Force Structure, Resources and Assessment, The Joint Staff
- The Director Program Analysis and Evaluation, Office Secretary of Defense

Table of Contents

Introduction and Workshop Summary.....	1
Workshop Annotated Report	9
Backup Slides.....	38
Working Group 1: Data Management.....	57
Working Group 2: Obtaining Data From Outside Sources.....	73
Working Group 3: Creating Data Internally	85
Working Group 4: Converting and Employing Data.....	99
Working Group 5: Technology Support	119
Synthesis Panel	145
Appendix A — Other Briefs	A-1
Keynote 1: DoD Data Enterprise	
John Osterholz	A-1
Keynote 2: DoD Analytic Agenda	
Charles Swett for Chris Lamb.....	A-5
Post Workshop Observations, Baselines and the Analytic Agenda	
Dr. Jim Stevens	A-17
Appendix B — Acronyms	B-1
Appendix C — Terms of Reference	C-1

Improving Defense Analysis through Better Data Practices

Introduction and Workshop Summary

BACKGROUND

The Military Operations Research Society sponsored a workshop on data best practices at the Institute for Defense Analyses 25-27 March 2003. Over 125 analysts and decision makers representing a cross-section of the Department of Defense (DoD), government agencies and the Ministry of Defence in Great Britain participated in the Workshop. This section summarizes discussions, observations and data best practices identified at the meeting.

PURPOSE

The purpose of the Workshop was to identify best practices for generating, collecting, developing, maintaining, disseminating and applying data and metadata to help support better analyses within the DoD. In addition, meeting organizers hoped to lay the groundwork for improved data policy and procedures within the DoD. Over the course of the three-day meeting, workshop attendees were exposed to a number of current data practices within the DoD and other communities. A plenary session set the stage for small working group discussions by providing strategic perspective and high-level insights into the issues. Following the plenary sessions, participants attended separate working group sessions each addressing a different aspect of the data enterprise. A separate technology working group assessed available tools and their potential for facilitating DoD data objectives. The workshop focused on the following questions:

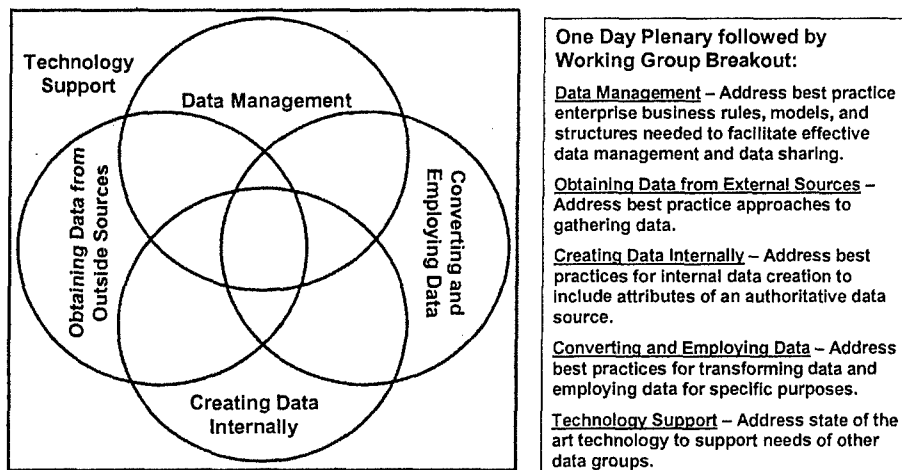
- What data content best encourages the development of analytic baselines to include methods of collecting and storing data, metadata, assumptions, scenarios contexts and other pertinent information?
- What cost effective methods and approaches best support verification and validation checks on data received, to include associated metadata in data repositories?
- What techniques are available to ensure data is appropriate for its intended uses be they extensions of existing applications or new applications?
- What effective and efficient methods are available for preparing data sets and associated metadata for use in specific studies and analyses to conclude the configuration of related data sets, delivering data according to criteria established by oversight bodies and archiving data and other study products for later reporting and application?
- What guidelines and methodologies are available for prioritizing data development, storage, management and dissemination activities?

While focusing on these questions, the workshop also identified organizations and activities within the DoD that are perceived to be authoritative data sources as well as suggested characteristics that could help define such sources.

STRUCTURE

Meeting Construct and Perspectives

Workshop Structure



Following an initial day of plenary presentations on data directions within the DoD and its analytic community, as well as discussions by data experts from other communities, the workshop broke into six working groups. A Synthesis Panel monitored the other working groups and constructed crosscutting insights validated by the other panels on the last day. The five other working groups took on specific aspects of data-related activity. The graphic above suggests the relationship between the subject matter addressed by these panels.

The Technology Support Panel overlapped all other activities, reflecting its charter to identify technological solutions to assist in the processes needed to support each of the other activities. The other four areas represent an arbitrary division of responsibility established by the workshop organizers to focus on more specific aspects of the data enterprise.

Data Management addressed enterprise business rules, models and structures needed to facilitate effective data management and sharing across DoD.

The Obtaining Data from External Sources (or Data Gathering) Working Group investigated how organizations acquire data from outside sources.

The Creating Data Internally (or Data Creation) Working Group addressed the methodologies and approaches for deriving data within an organization for specific uses, as well as identified expected attributes of authoritative data sources.

The last group, Converting and Employing Data, changed their name at the event from Transforming and Employing Data, to avoid confusion with transformation terminology currently used to describe DoD organizational goals. This group looked at procedures and approaches for converting available data (often through aggregation or disaggregation) and employing it for specific purposes, usually within models and simulations to achieve analytic results.

PLENARY SESSION

The Workshop's plenary speakers addressed DoD data goals. Although there was broad consistency across presentations, there were also subtle differences. John Osterholz, OASD(C3I), in his talk on Enterprise Data Strategy, formulated two broad goals supported by several subordinate objectives. He emphasized that his office is focusing on the "availability" goal (and its subordinate objectives of visible, accessible, and institutionalized). This is consistent with the ASD(C3I) focus on providing the community with a Global Information Grid that provides very wideband connectivity among key users. In addition, his office has the goal of enhancing data "use" (subsuming the objectives of providing data that is understandable, trusted, interoperable and responsive to user needs). (See Appendix A.)

Charles Swett, who appeared for Chris Lamb, OSD(Policy), identified and emphasized several different goals and objectives. He placed particular emphasis on the need for data that are "rapidly available," with consideration for the need to be transparent, accurate, standardized, sufficiently detailed and synchronized with PPBS activities. He also observed that "As DoD transforms, the analytic system must transform" and emphasized several attributes that represent a transformed analytic system: higher quality analyses; better synchronized with key institutional processes (e.g., PPBS); *more flexible and responsive*; and, with an enhanced ability to cope with unanticipated needs. In addition, he stressed that the transformed systems must be able to deal rapidly with a wider range of variation and/or future uncertainties. (See Appendix A.)

Tony Simon, OASD(C3I), presented two similar, but slightly different perspectives on relevant goals. Initially he observed that the goal was to "ensure that all data are visible, available, and useable when needed and where needed." He subsequently concluded his remarks by citing the goal to "make data visible, accessible and understandable across DoD and beyond."

Jim Stevens, in his presentation on the new DoD Data Directive, summarized his presentation by identifying the goal of the directive as being the production of "visible, accessible and traceable data ready to support analysis."

Overall, the data enterprise presents a multi-attribute utility problem with some perceived differences in the components and emphasis in the objective function. These issues will need to be clarified as the community continues to address this problem area.

These speakers were followed by a panel discussion to provide service and joint staff perspectives. Panel members included Dr. Jackie Henningsen, FS, Director of the Air Force Studies and Analyses Agency; Mr. Vern Bettencourt, FS, Technical Advisor to the Army Deputy Chief of Staff for Operations and Plans; Mr. Greg Melcher, Deputy Director for Assessments (N-81); Dr. George Akst, Deputy Director Studies and Analysis Division, Marine Corps Combat Development Command; and Mr. Pete Byrne and Mr Kevin Kelley, representing the analytic functions of J-8 and J-4.

The panel in general reinforced primary themes introduced by the plenary speakers and also noted additional actions that must be undertaken to transform the DoD data enterprise and the analytic system. These include the ability to perform capabilities-based planning, the need for joint operational concepts for future operations, the requirement for a wide array of approved tools (particularly those that are able to treat the challenging problems associated with Information Operations (IO), Weapons of Mass Destruction (WMD) and C4ISR), the collectional use of real-world data from ongoing operations, and the need for enhanced analytic resources (e.g., adequately educated and trained people). Challenges include the Department's current culture, lack of standardized data management approaches and systems, security issues, and lack of consistency in substance, format and processes across Services and Agencies.

These data challenges are not unique to the DoD. Expert presenters from the National Oceanographic and Atmospheric Agency (NOAA), the National Science Foundation, the National Institute for Health and the State Department suggested that even well defined and long standing areas of endeavor require significant improvement in their data processes.

BEST PRACTICES

The workshop identified a number of trends and methodologies seen as "best practices," not only for DoD analytic and study data but worthy of emulation in other arenas. In particular the trend toward funding VV&A of models and data as part of program development rather than an after-the-fact add-on is a policy change that is already contributing to improved credibility of tools and the data required to use them. Another best practice has been the government's desire to include community input into the creation of standards and new practices. The Data Workshop was one beneficiary of this attitude. In the technology arena, the use of existing COTS and GOTS tools enable current processes and help the community in self organizing to meet important objectives in the data arena. A process improvement in the intelligence arena worthy of emulation is the movement away from processes focused on the producers of intelligence to methods focused on intelligence users. This same shift in emphasis in the analytic data area is enabling the creation of new tools and processes to focus on the analyst and the decision maker, rather than the individual producer of data and will help in the long run to establish appropriate data methodologies and practices of greater benefit to the community.

The workshop also reinforced the use of Metadata as a key to improvement in data management. An excellent tool in this area is the DoD's Discovery Metadata Standard that is already available to help facilitate data management and exchange. Implementation will make data more accessible to the using community. Infrastructure and services that already exist, such as Centric Enterprise Services, Horizontal Fusion, shared spaces and metadata catalogues and registries will all assist in moving the Department's data enterprise to meet the objectives set for the community. The Army's Materiel Systems Analysis Activity (AMSAA) and OSD's Joint Data Support (JDS) have also developed and implemented process and product enhancements including a number of useful data maintenance and data management concepts.

In addition to current best practices, speakers and other workshop participants identified a number of other candidate best practices that could be adopted for DoD use. For example, in the policy arena, the Department could provide incentives for data sharing to encourage culture change in the Department. The benefits of having good data shared across service, command, community and activity lines far outweigh the costs to individual organizations of providing such data. The Department could develop a range of carrot and stick alternatives to move the community toward facilitated data sharing. Using available systems and technologies to organize around and empower the data users vice the data generators would go a long way to improving practices in this area.

With respect to organizations, a central authority and repository for all DoD data was seen as not only not possible due to the literally thousands of activities that generate, convert and use data for different purposes, but not desirable as well. Instead, the national library system was posed as a model for what could be a more effective data enterprise. A central data curator in conjunction with a senior joint oversight or steering group could identify a unilateral structure for core and common data requirements, and assign organizational responsibility to the appropriate organizations for the maintenance of these broadly required items. Such a joint DoD data committee would be made up of representatives of the services, OSD principals and Joint staff assisted by appropriate technical expertise from the Federally Funded Research and Development Centers (FFRDC), academia and industry. The curator and committee would work top-level issues, such as establishing definitions, metadata standards, a controlled data vocabulary and identifying new technologies.

In the tool area, the Department could consider developing standard algorithms and methodologies to ensure modeling and simulation data are used in a manner that is consistent with its original definition and context. Standardized language and metadata would facilitate the appropriate use of data. A number of other process improvements could also enhance DoD data practices. These include the incorporation of library science techniques as well as processes to improve visibility and accessibility of data resources. Other improvements for consideration include the establishment of formal mechanisms to simplify accreditation and adaptation of accredited data for other analytic or operational purposes. Likewise, sponsoring activities that improve dialogue between those who develop data and those who use data for analytic or other purposes would set the stage for improved management and application of data across communities. Such activities include MORS workshops but could also incorporate a range of initiatives designed to connect the data production and data application communities together more effectively.

Other process best practices could include implementation of formal data audit trails for data organizations gathering, creating and using data, in order to guard against misapplications and help inform future users of the data about its character and limitations. Peer reviews, collaborative working groups, metadata standardization and other techniques could all assist in this activity.

The workshop noted that standardized processes are most appropriate and effective within organizations and not across organizations. However, the community needs a forum for sharing these processes so that organizations can select from an array of processes that work to ensure internal standardization occurs. The Department can facilitate these exchanges as well as put in place useful business rules that better connect organizations and facilitate data exchange.

Finally, best practice products include a controlled data vocabulary and a unilateral structure of common and core data as noted earlier in the organizational best practice section.

WORKSHOP SUMMARY

Workshop participants agreed that the data problem is enormous ... and increasing in both size and complexity. The workshop served to showcase several significant initial steps that the community has taken to address the problem. First, there was broad agreement that the new directives (DoDD 8260.1) and instructions (DoDI 8260.2) provide a useful foundation for attacking an important segment of the problem. There was anticipation that the proposed Analytical Baselines will serve to enhance the speed and quality of future strategic analyses. In addition, there is hope that the Joint Analytic Data Management Steering Committee (JADMSC) will provide an effective forum for overseeing the creation and evolution of these Analytical Baselines. Further, the framework promulgated by OASD(C3I) to conceptualize the problem (i.e., enterprise, community of interest, private data) appears useful along with the new tools and standards that they are developing (e.g., the DoD Discovery Metadata Standard (DDMS)).

However, in order to make additional substantive improvements, it is vital that the military analysis community take several challenging steps. The most important is to start now to transform the culture of data from one of hoarding to one of sharing. To do so, steps are needed to dispel the fears that permeate the community (e.g., fears of misuse, misunderstanding and adverse consequences). This initiative must be undertaken and sustained at the highest levels of leadership. Second, the people issues must be addressed. This entails educating and training the users and providers of data – as well as the decision makers. Finally, the key technical issue to be worked is that of metadata. This theme was mentioned by most of the plenary speakers and it constituted the baseline recommendation of the working group on data management. The problem is difficult, but the community must begin to address it seriously immediately.

A scorecard to summarize the accomplishments of the event and the residual challenges that confront the community helps put the workshop in perspective. In abbreviated form, the workshop sought to achieve four objectives: 1) review current data processes; 2) identify best practices and standards; 3) identify and assess potential sources of data; and, 4) educate the community on the new data directive. The workshop made excellent progress on two of those

objectives (review current data processes and educate the community on the new data directive), and made a good start toward achieving the remaining two objectives (identify best practices and standards; identify and assess potential sources of data). Given the highly complex, multi-dimensional problem of interest, these were major accomplishments.

However, this constitutes a beginning, not an end. Since we are now ready to begin, the workshop pointed toward three major initiatives that the community should undertake. First, there is a need to reach out to a broader community to address the data problem. This includes enlisting the aid of the inter-agency and alliance/coalition communities as well as adapting the approaches and lessons of the library science community. Second, there is a need to generate a plan of action and milestones to focus and mobilize the analysis-data community. Such a product would help coordinate diverse actions and provide the discipline needed to ensure steady progress. Finally, the community must recognize that the complexity and size of the problem is such that a single workshop is inadequate. MORS should work with its Sponsors to identify and initiate follow on activities to address many of the critical issues identified during this workshop.

WORKSHOP PRODUCTS

The next section of this document consists of an annotated Workshop Report Briefing. Of note, a substantial portion of this briefing was extracted and briefed to the US Department of Defense Joint Analytic Data Management IPT and Steering Committee in July 2003, at their request. In particular the IPT used the workshop information to reinforce lessons and issues being addressed by those bodies as well as to help individual members of the Steering Group to think through potential incentives to facilitate sharing data. Two additional slides were created for those occasions that now appear as the first two backup slides in the annotated Workshop Report Briefing.

Following the Workshop Report are sections comprised of out-briefs for each of the five Working Groups and the Synthesis Group. These observations formed the basis for the Workshop Report Briefing.

Appendices contain the two Keynote addresses and post-workshop observations by a key participant; a list of acronyms; and the workshop Terms of Reference.

A summary report appeared in *PHALANX* Volume 36 Number 3, September 2003.

**Improving Defense Analysis
Through Better Data Practices**

25-27 March 2003
Alexandria, VA



Annotated Workshop Report Briefing

12 June 2003
Chairs: Tom Allen
Jim Bexfield, FS

Overview

- **Background**
- **Meeting Construct and Perspectives**
- **Working Group Issues and Observations**
- **Best Practices**
- **Workshop Observations**
- **Summary**

Backup

- **Definitions**
- **Data Sources**
- **Other**



This report consists of five primary sections and a backup reference section.

The first section provides background information for the workshop sponsored by the Military Operations Research Society (MORS) at the Institute for Defense Analyses (IDA) on 25-27 March 2003. The purpose of the workshop was to improve analysis through better data practices.

The second section outlines how the workshop was constructed and summarizes perspectives on the nature of the problem developed from the remarks of the plenary speakers.

The third section provides observations by speakers and individual working groups as developed by Synthesis Panel members and reported by the various working groups that made up the workshop.

The fourth section provides a list of current and potential future data best practices for DoD consideration as presented during working group presentations and discussions by the various panels.

The final section summarizes key recommendations and conclusions.

Among other items, the backup section offers a start to data definitions as well as lists some current data sources identified by participants and workshop organizers over the course of the workshop.

Workshop Objectives

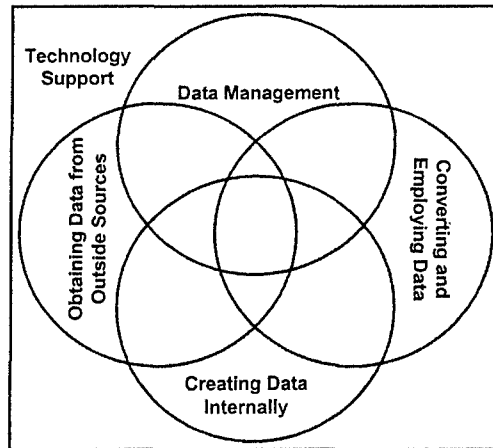
1. Review current DoD processes for developing, managing, using and sharing data in support of military analyses to include defining common terms
2. Identify best practices and standards within DoD and the broader analytic community for generating, collecting, developing, maintaining and disseminating data; suggest how applicable best practices could be implemented in DoD
 - Review expected impact of technology on practices and processes
 - Note issues and problems that serve as barriers to implementing best practices
3. Identify and assess potential sources of data needed for a wide variety of analyses
 - Provide a current listing of authoritative data sources in the Department
4. Educate community on the new Data Directive, to include plans for implementation (data repository, access to the data, etc.)



The workshop was constructed to meet four major objectives. The first was to review current DoD processes in various aspects of the DoD data enterprise. The second was to identify the best practices and standards for data, both within the Department and across the wider academic, government and industrial communities. The third goal was to identify authoritative data sources for the Department. Finally, the workshop was intended to provide an opportunity to educate the community on data developments within the Department as well as address the contents of the newly released Data Directive.

In accomplishing these goals, the workshop helped to clarify the nature of the problems and issues associated with the data enterprise in DoD through a series of general presentations on the first day of the workshop and then through presentations and discussions by working groups designated to address various parts of the data enterprise over the remaining period of the event. Second, it sought to capture the state-of-the-practice in data practices, based on the workshop activities. Finally, it derived key findings and recommendations by integrating the information received and by the use of a dedicated Synthesis Panel, established to look across the results of the individual panels to find synergistic and general insights.

Workshop Structure



One Day Plenary followed by Working Group Breakout:

Data Management – Address best practice enterprise business rules, models and structures needed to facilitate effective data management and data sharing.

Obtaining Data from External Sources – Address best practice approaches to gathering data.

Creating Data Internally – Address best practices for internal data creation to include attributes of an authoritative data source.

Converting and Employing Data – Address best practices for transforming data and employing data for specific purposes.

Technology Support – Address state of the art technology to support needs of other data groups.



In addition to the Synthesis Panel, five other Working Groups (WGs) participated in the workshop. The above graphic suggests a relationship between the subject matter addressed by the groups. The Technology Support WG overlapped all other activities, reflecting its charter to identify technological solutions to assist in the processes needed to support each of the other activities. The other four areas represent an arbitrary division of responsibility established by the workshop organizers to focus on more specific aspects of the data enterprise.

Data Management addressed enterprise business rules, models and structures needed to facilitate effective data management and sharing across DoD.

The Obtaining Data from External Sources (or Data Gathering) Working Group investigated how organizations acquire data from outside sources.

The Creating Data Internally (or Data Creation) Working Group addressed the methodologies and approaches for deriving data within an organization for specific uses, as well as identified expected attributes of authoritative data sources.

The last group, Converting and Employing Data, changed their name from Transforming and Employing Data, to avoid confusion with the use of the transformation term to describe DoD organizational goals. This group looked at procedures and approaches for converting available data (often through aggregation or disaggregation) and employing it for specific purposes, usually within models and simulations to achieve analytic results.

Key Personnel

General Chair	Tom Allen
Co-Chair	Jim Bexfield, FS
Plenary Facilitator	Scott Simpkins
Data Management WG	Simone Youngblood, Jim Stevens
Data Gathering WG	Jack Sheehan, Philippe Loustaunau
Data Creation WG	Bryan Paris, Dennis Leedom
Data Employment WG	Clay Bowen, Ernie Boehner, Rudy Pabon
Data Technology WG	Bob Might, Ron Smits, Jim Richardson
Synthesis WG	Stu Starr, FS, Kirk Yost
Admin Coordination	Brian Engler, Natalie Kelly
MORS Bulldog	Bill Dunn
OSD Rep	Jim Bexfield, FS
Joint Staff Rep	Bob Orlov
Army Rep	Paul Deitz
Navy Rep	Herb Cupo
Air Force Rep	Roy Reiss
Marine Corps Rep	Darren Whaley



Putting on a workshop requires a motivated and dedicated team. Fortunately for the chair, such a team was available and they are singularly responsible for the success of the meeting. In particular, the workshop would not have happened without the early engagement and clear focus provided by Jim Bexfield, a Fellow of the Society and the Office of the Secretary of Defense Sponsor's Representative for Mr. Eric Coulter from Program Analysis and Evaluation (PA&E).

Working Group Scope, Objectives

- **Scope**

- Focus on data to support *strategic analyses*
- Consideration given to data to support analyses for acquisition, OT&E, operations

- **Objectives**

- Clarify problems by conducting internal panel discussions
- Capture the state-of-the-practice based on the presentations at the Plenary
- Derive key findings and recommendations (including best practices) from the results



In order to better scope and manage the content of the discussions in the working groups, the workshop focused primarily on data to support strategic analyses. However, where appropriate, consideration was also given to the issues associated with data to support analyses of other defense issues (e.g., system acquisition, Operational Test and Evaluation (OT&E), operations) since most participants agreed that the same general insights would accompany a more detailed look at each of these areas. In addition, many of the insights apply to data in support of other activities to include operations, experimentation, test, finance and administration.

Each working group and the Synthesis Panel pursued three supporting objectives primarily through expert presentations and group discussions. First, the groups sought to clarify the nature of the data problem and identified important issues from the perspective of their focus area. Second, the groups sought to capture the current state of practice for those dimensions of the data problem relating to their area, based on the discussions and presentations. Finally, each group derived key findings and recommendations by incorporating observations and inputs from the participants in their working group. The Synthesis Panel then integrated across working groups to produce a comprehensive presentation that addressed the full dimensions of the data process. The inputs from each of these groups formed the basis for this brief.

Not a New Problem

"The government are very keen on amassing statistics. They collect them, add them, raise them to the n-th power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of these figures comes in the first instance from the village watchman, *who just puts down what he damn pleases.*"

Comment of an English judge on the subject of Indian statistics

"It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."

Sir Arthur Conan Doyle

"Theory without data = philosophy; data without theory = noise"

Anonymous

"Without data we are nothing!"

Walt LaBerge, SIMTECH 1997



As a cautionary tale, the Synthesis Panel and others identified several quotes that served to dramatize several dimensions of the analysis - data problem. The first one above from Sir Josiah Stamp underscores the fact that extensive processing of data should not obscure the fact that the initial information itself may lack quality or credibility.

The other quotes bring us up to the current day, underscoring both the importance of data and the necessity for developing a common approach to its development.

Plenary: Views of Data Goals

- **John Osterholz, Enterprise Data Strategy Goals**
 - AVAILABLE: visible, accessible, institutionalized
 - Use: understandable, trusted, interoperable, responsive to user needs (requires a culture change)
- **Charles Swett (for Chris Lamb), DoD Analytic Agenda**
 - Transparent, RAPIDLY available, accurate, standardized, sufficiently detailed, synchronized with PPBS
- **Tony Simon, DoD Data Management**
 - Ensure that all data are visible, available, useable when and where needed
 - Make data visible, accessible and understandable across DoD and beyond
- **Jim Stevens, DoD Data Directive**
 - Visible, accessible and traceable data ready to support analysis



All four speakers during the initial plenary session addressed the subject of DoD data goals. Although there was broad consistency across presentations, there were also subtle differences worth noting.

John Osterholz, OASD(C3I), in his talk on Enterprise Data Strategy, formulated two broad goals supported by several subordinate objectives. He emphasized that his office is focusing on the “availability” goal (and its subordinate objectives of visible, accessible and institutionalized). This is consistent with the ASD(C3I) focus on providing the community with a Global Information Grid that provides very wideband connectivity among key users. In addition, his office has the goal of enhancing data “use” (subsuming the objectives of providing data that is understandable, trusted, interoperable and responsive to user needs).

Charles Swett, who appeared for Chris Lamb, OSD(Policy), identified and emphasized several different goals and objectives. He placed particular emphasis on the need for data that are “rapidly available,” with consideration for the need to be transparent, accurate, standardized, sufficiently detailed, and synchronized with PPBS activities.

Tony Simon, OASD(C3I), presented two similar, but slightly different perspectives on relevant goals. Initially he observed that the goal was to “ensure that all data are visible, available and useable when needed and where needed.” He subsequently concluded his remarks by citing the goal to “make data visible, accessible and understandable across DoD and beyond.” Although “available” is related to “accessible” and “useable” is related to “understandable,” those terms are not truly congruent.

Jim Stevens, in his presentation on the new DoD Data Directive, concluded his presentation by identifying the goal of “visible, accessible and traceable data ready to support analysis.”

Overall, the data enterprise presents a multi-attribute utility problem with some perceived differences in the components and emphasis in the objective function. These issues will need to be clarified as the community addresses this problem area.

Plenary Major Points

- **“As DoD transforms, the *analytic system* must transform”**
 - Perform analyses that are higher quality, better synchronized, more flexible, responsive, able to cope with unanticipated needs
 - Deal with a wider range of variation/future uncertainties
- **Transforming the *joint data* enterprise is just one element of the analytical transformation; it also requires**
 - Capabilities- Based Planning
 - Joint Operational Concepts
 - Wider array of approved tools (e.g., deal with IO, C4ISR)
 - Enhanced analytic resources
- **Challenges include current culture, lack of standardized data management approaches and systems, security issues, and lack of consistency in substance, format and processes across Services and Agencies**



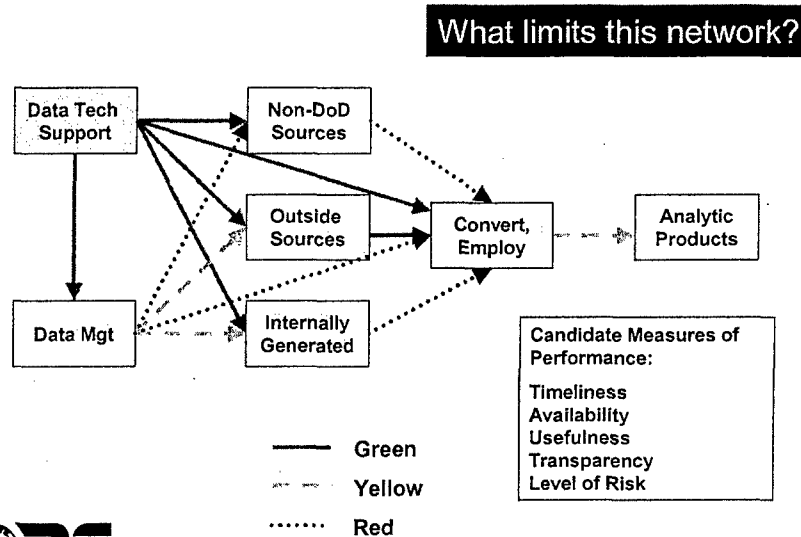
These four speakers were followed by a panel discussion to provide service and joint staff perspectives. The panel included Dr Jackie Henningsen, FS, Director of the Air Force Studies and Analyses Agency; Mr. Vern Bettencourt, FS, Technical Advisor to the Army Deputy Chief of Staff for Operations and Plans; Mr. Greg Melcher, Deputy Director for Assessments (N-81); Dr. George Akst, Deputy Director for Analysis, Marine Corps Concepts and Doctrine Center; and Mr. Pete Byrne and Mr. Kevin Kelley, representing the analytic functions of J-8 and J-4.

Charles Swett observed, that “As DoD transforms, the analytic system must transform” and emphasized several attributes that represent a transformed analytic system: higher quality analyses, better synchronized with key institutional processes (e.g., PPBS), *more flexible and responsive*, and with an enhanced ability to cope with unanticipated needs. In addition, he stressed that the transformed systems must be able to deal rapidly with a wider range of variation and/or future uncertainties.

Some of these themes were reinforced by others who also noted additional actions that must be undertaken to transform the DoD data enterprise and the analytic system. These include the ability to perform capabilities-based planning, the need for joint operational concepts for future operations, the requirement for a wide array of approved tools (particularly those that are able to treat the challenging problems associated with Information Operations (IO), Weapons of Mass Destruction (WMD) and C4ISR), and the need for enhanced analytic resources (e.g., adequately educated and trained people).

Data challenges are not unique to DoD. Expert presenters from the National Oceanographic and Atmospheric Agency (NOAA), the National Science Foundation, the National Institute for Health and the State Department suggested that even in well defined and long standing areas of endeavor, much improvement in data processes is required. Specific challenges raised by the first day’s speakers are noted in the last bullet of this slide.

The Working Groups: An Assessment



The Synthesis Group rearranged the working group Venn Diagram to show supporting dependencies. Based on their deliberations, they provided a grading criteria that might be applied to DoD analytic data processes currently in use. This slide represents their overall assessment of the key nodes and links in the analysis-data framework, employing the familiar “stop light” scale of red (poor), amber (fair) and green (good). The assessment reveals that the biggest problems are not so much in the “nodes” (e.g., availability of technology, data management, etc.), but in the links between the nodes. While the evaluation of any node (with respect to measures of performance such as timeliness and availability) could be debated, there was broad agreement that many of the links are “red” (i.e., not functioning).

A notable exception to the link assessment concerns those connecting data technology solutions to the other areas. In general, technology is available or emerging, and, in some cases, is already in place. The rest of the links, however, are not working effectively, particularly those between data providing nodes and the converting/employing data node.

The normal bureaucratic solution to such situations has an interesting analogue in physics. The application of heat makes a physical substance flow more easily. The normal day-to-day arrangements for DoD data are characterized by low flow and high viscosity (or “resistance”) that is often overcome only by heat being applied by the decision maker.

This assessment does not prejudice the outcome of new initiatives such as the DoD regulations on data management and analytic baselines. It does, however, point out that the problem is largely *between*, and not *within*, organizations.

The following charts summarize the observations of the five working groups. For each working group the charts attempt to summarize the perceived nature of the problem and then provide key observations.

Data Management

- **Data Management issue is vast, uneven in maturity across DoD, and lacks a comprehensive working process to meet the performance measures identified by the plenary speakers**
- **Observations:**
 - Most effective approach would be a decentralized (networked) repository structure
 - Requires definition of a disciplined framework (e.g. JSF example)
 - Suggests need for a joint DoD committee to establish and manage
 - The business model surrounding data needs improvement
 - Existing situation has clear disincentives for data sharing (e.g., time to maintain and productize, cost, negative impacts of bad outcomes)
 - Options for improving incentives for good data management are available (financial, report card, measures, considerations) but rarely implemented
 - Business model should incorporate processes and procedures for re-use and risk management
 - Key improvement would be the extension of policy over and enforced use of Metadata



There is a significant amount of process and structure in place within DoD to support data management, but the scope of the data management requirement is vast, even within the restricted analytic modeling realm, and the state of the process and infrastructure is uneven across the Services and other data holders, including DoD and the non-DoD government communities. After reviewing Service and DMSO data management processes and procedures, the workshop concluded that a comprehensive departmental process that might lead to the achievement of the performance goals and measures provided by the plenary speakers is lacking.

Consequently, workshop participants identified three major observations. First, a more decentralized (networked) repository structure seems to offer the most potential for meeting DoD data management goals. In order to build such a structure, the Department needs to define and implement a disciplined framework. As an illustration, the Joint Strike Fighter (JSF) approach provides a useful model of such a framework. A high level joint DoD management committee could establish and oversee this framework with representatives from the Services, Joint Agencies and be supported by appropriate technical expertise (i.e., FFRDCs, academia)

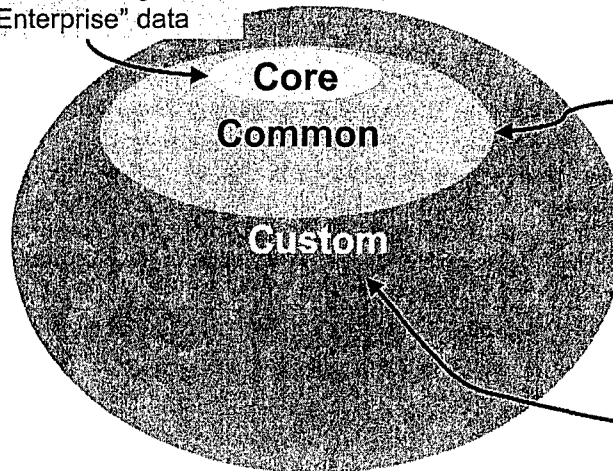
Second, the business model surrounding data needs improvement. Even though “lines of communication” for data flow exist in many areas, the existing situation not only lacks clear incentives, it is characterized by serious disincentives. As examples of the latter, the current process requires significant investments of time to maintain and productize data, and incurs substantial costs. Decision makers can choose from a variety of options to improve incentives (e.g., provide additional resources for following best data practices). In addition, establishing processes and procedures for re-use and risk management would allow straightforward and consistent appraisals of data-related risks to be conveyed to decision makers, thereby affording some flexibility in data management.

Finally, the single most important step would be to recognize metadata as the key enabler to effective data management. Extending and enforcing metadata policy would be a major positive step forward.

Observations

Obtaining Data from Outside Sources

Without this there is no scenario; analogous to "Enterprise" data



Most scenarios use some of this; similar to "Community of Interest" data

Somewhere someone needs some of this; similar to "Private" data

nors

In considering the topic of obtaining data from external sources, the workshop took a lead from the plenary presentations and other work in this area to focus on two triads: 1) Core, Common, Custom; and, 2) Enterprise, Communities of Interest, Private. The first seemed to refer more to the data sets themselves while the second formed around the data producers-users, although the workshop did not achieve full agreement on the definitions. Working definitions for these and other terms is included in the backup material to this brief.

Defense data gathering (in fact, defense analysis) can be seen as a two-fold problem from the data gathering perspective: normal or peacetime support to DoD (e.g., allocation of resources, training, deliberate) and response in crisis (e.g., distribution of available resources, haste). The overall structure for analysis and data can be broken into mission-to-task statements rolling down the hierarchy, with succeeding missions and tasks break-outs becoming ever more detailed. Thus, task statements can provide the direction and scope of needed data.

While workshop participants agreed that there are relationships between the two triads, they offered different views of the specifics of these relationships.

The community will need to conduct further examination of the potential relationships between these triads as well as produce more careful definitions of the triads themselves. Also, while it is reasonably clear that the mission-to-task structure can roll down the hierarchy, it is not clear that the resulting data needs can roll back up the other direction. The main point is that some data development and requirements can be met in advance by a single data developer for all data users, other needs can be met by a group of data developers working together to provide data for a range of scenarios, while a final set of data will need to be developed at the time of a study or analysis and will be unique for the specific use.

Obtaining Data

- **Most current data needs are defined from mission-task statements as a function of level of “war”**
 - Qualified by distinction between normal operations (peace) and response operations (crises).
- **Observations (five elements for success noted):**
 1. Requirements Engineering with Users
 - The more complete the better
 2. Social Engineering with Stakeholders
 - Must enlist voluntary cooperation
 3. Technical Engineering with R&D Community
 - System/technical architecture should be Technology Readiness Level 8
 4. Information Engineering to meet data user needs
 - Demand pull (users pull products, products pull tools)
 5. Inventory Management
 - Stock components that can be tailored for custom assembly



In addition to breaking down data needs from mission-task statements and its two-fold applications (e.g, normal or peacetime to address allocation of resources, training, deliberate; and, response or crisis to address distribution of available resources in an accelerated decision environment) workshop participants evaluated a number of real life data gathering situations. From these, five elements of success seemed to be associated with situations where the needed data was collected efficiently and with a maximum of fidelity.

Elements of success included early engagement by the data users with the providing community, frequent communications between all elements of the enterprise for which the data was being developed, understanding of the technical issues and the range of potential results associated with lower technical readiness levels, and the application of facilitating information management rules. Also considered necessary are procedures and the ability to create the right levels of core and common data for use by other organizations, as well as to hold the appropriate data that will have to be custom created for specific applications. Successful data enterprises worked each of these areas and had in place procedures, methods and approaches to ensure all elements of the enterprise worked effectively together to minimize friction and maximize the production of useful data for the analytic purpose for which it was intended.

Creating Data Internally

- **Many obstacles and barriers to doing this efficiently across the department**
 - State of art of performance modeling (individual; organizational)
 - Lack of unified theory of war
 - Data aggregation - nonlinear effects
 - Large community of data providers – presents a large information management problem
- **Observations**
 - There are several promising approaches, such as those applied by AMSAA and JDS as well as data audit trail/pedigree processes
 - Need a more comprehensive approach to establishment and use of data standards
 - Nomenclature/operational definitions/categories
 - Resolution/fidelity
 - Format/tagging/metadata
 - Need to link modeling paradigms to data standards
 - Need to develop a more universal process for data requests across DoD organizations



The workshop noted several barriers to efficient and effective data creation. Addressing the problem required identifying the data content associated with various types of data produced to support joint studies, experiments and wargames; listing potential sources for this data; and, understanding and evaluating how data are developed by various sources.

Participants addressed three categories of data domains: 1) socio-cognitive (extremely difficult to generate data); 2) information (challenging); and, 3) physical (relatively easy). Focusing on the socio-cognitive and information domains, the workshop identified several barriers and obstacles to data generation. The lack of a unified theory of war makes it difficult to achieve consensus on an appropriate set of performance models which leads to an inability to develop a consistent set of data requirements. There are non-linear effects associated with aggregating or scaling data (e.g., individual behavior does not scale to group behavior, which in turn does not directly scale or aggregate into organizational behavior). The large number of data providers with unique data formats, combined with the need to use data from multiple exercises, experiments, and tests in order to build a robust data set, presents a significant data management and data mining problem.

There are several promising approaches to address problems in this area. In particular, the data archiving, distribution and control methods practiced by Army's AMSAA and OSD's JDS represent concepts that deserve further investigation and adoption. In addition the data audit trail and pedigree process of the Air Force's Aeronautical Systems Center provides an excellent model for controlling and tracing internally created data.

Other observations are that the community needs to develop and publish a common set of data standards, including operational definitions, a common nomenclature, and data categories. Because one model's output is frequently another model's input, there is also a need to develop appropriate sets of processes/data transformation standards to enable this linkage. Finally, community development of a more common process for making, reviewing and sharing data among the different data providers would offer a significant improvement to this aspect of the DoD data enterprise.

Converting and Employing Data

- **Several issue areas: data conversion, data organization and storage, data security and data sharing**
- **Wide variety of approaches to internal data management and generation**
 - Trust issues with sharing (misuse and risk to provider)
- **Observations**
 - Most effective results seem to come from organizing within provider organizations rather than attempting to standardize across organizations
 - A useful contribution at the OSD level would be to develop an effective external structure for sharing
 - Community must aggressively work the trust issue; DODD/DODI and analytical baselines seem the best place to start



In the data conversion and employment area, the workshop noted several issues. Among these were data conversion, data organization and storage, data security, and data sharing.

Several definitions for types of data conversions are included in the backup material.

With respect to data storage and organization, the workshop suggested that a single approach to data organization, which the DoD attempted in the mid-1990's, is inappropriate. However, participants noted that each providing and using organization must organize *internally* for data storage and sharing before it can effectively share data externally. Different organizations can (and probably should) use different schemes, but each organization needs to develop and implement a clearly articulated approach. Once such an internal data storage structure exists, it will be possible to share data across organizations using some minimal external structure, using utilities and standards that convert data for specific needs. Such a process would be facilitated by the application of clear metadata standards and approaches.

While the physical approaches to storage and sharing seem straightforward, the trust aspect of sharing is much more difficult. In fact, a major insight of the workshop is that there are few technological barriers to sharing, but the "trust" issue is what precludes most data sharing in DoD. These can be summarized as a variety of "fears:" fear of misuse due to other users not understanding the data; fear of being held to available data even if such data is incomplete or nonrepresentative of broad applications and particularly when used for future forecasts; and, fear that data supplied will be used to undermine the very systems and force structures the provider is attempting to understand before a full series of insights and applications can be developed. The *trust* issue is the most critical one to overcome.

Data Technology Support

- **Issues spring from several sources**
 - Lack of understanding of the supported analytical processes
 - Data business processes that begin with the available supporting technology rather than the supported analysis
 - Disjointed data technical processes
 - Data tools that lead rather than follow the analytic process
- **Observations**
 - The application of standard system engineering theory in the use of existing technologies or development of new technologies can simplify and improve the process
 - A good first step would be to standardize the generation of metadata
 - Would also be helpful to further define business processes
 - Organizations should first define their desired analytic and data management processes before applying existing technical process or tools, or developing new ones

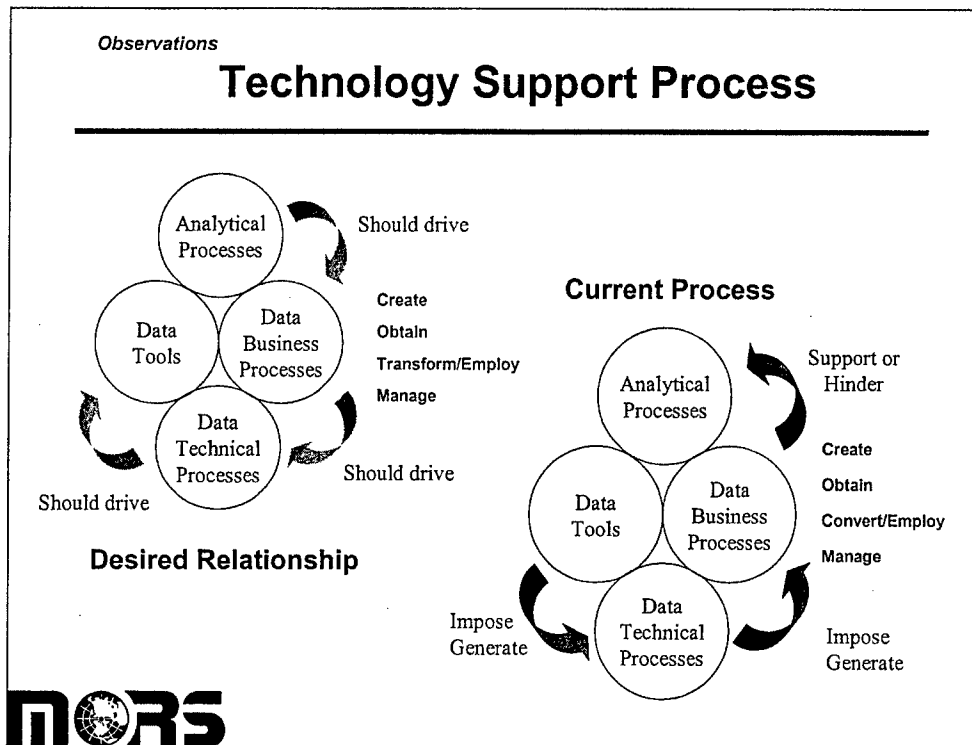


Workshop participants determined that technology support issues can be related to the processes involved in the data enterprise (i.e., analytical processes, data business processes, and data technical processes), the proper sequence among these processes, and the supporting tools. Proper ordering suggests analytical processes should drive data business processes. Once these are established, the data business processes should drive data technical processes. Finally, data technical processes should drive data tools. An example of this sequence is the generation of survivability data. First the survivability analysis is articulated (analytical process is defined), threat data is created by this analysis (the generation of a data business process), and then ways to rapidly create the supporting data from engineering models is determined (the application of data technical processes). Once the analytic, business and technical processes are in hand, the MATLAB/SimuLINK tool can be tailored and applied (showing how a data tool can then be created to support the other processes).

Since often this system engineering approach is turned on its head (tools drive technology which in turn drive business processes to constrain the analyses), participants observed that the correct application of systems engineering theory is the use of existing technologies or the development and use of new technologies would improve the overall application of technology to the data enterprise.

Of many considerations, standardizing the generation of metadata deserves special emphasis because it significantly supports virtually all the business processes (and in particular the useful sharing of data).

In addition, it would be helpful to redefine or refine the business processes to avoid overlaps and focus on the end product. Finally, organizations should take the time to define their desired analytic and data management processes before applying technical solutions to store and manage their data.



This chart illustrates the discussion of the previous slide. The systems engineering desired order of steps is as shown on the left: identify analytic processes to be supported; convert analytic processes to supporting business process requirements; identify and define technical processes that support business process requirements; convert technical processes to technology requirements; and, based upon technology requirements, develop the technologies and evaluation tools to assess existing technological and non-technological solutions and develop requirements for new solutions.

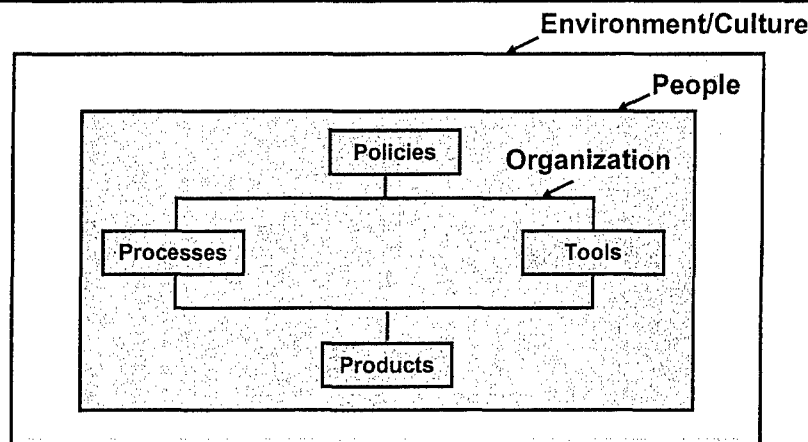
Analytical processes should drive data business processes, i.e., manage, obtain, create, convert.

Data business processes should drive technical processes and the technical processes should drive the data tools.

Unfortunately, in many cases, the availability of a data tool will suggest that it be used and therefore impose a technical process. Or when new tools are developed they become the driver to formulating new technical processes. These technical processes in turn impose data business restrictions and processes that may or may not be useful in the analytical process that they support. In many cases, they may actually hinder the analytical process.

The result is the illustration on the right of the chart, which all too often represents the current state of practice.

Framing the Overall Observations



The functions of an organization are dependent on the quality and experience of its people who are in turn influenced by their culture and environment. We need to address all these in order to ensure a comprehensive solution.



The accompanying figure provides a business process re-engineering perspective of the data - analysis enterprise. The workshop concluded that if enhanced data practices are to truly improve defense analysis, we must consistently address *all* of these factors. The backdrop for these factors is set by the *cultures* of the many communities that must participate in the production and use of data for analysis. It was recognized that in many cases DoD would rely on the data provided by non-traditional partners (e.g., NGOs). Thus, we must be cognizant of the cultures of the other participants and flexible in our interactions with them.

Second, *people* are critical components of the data enterprise. This implies the need to provide critical Education and Training (E&T) for the providers and users of data as well as the decision maker. Third, we may have to change *organizational* relationships to facilitate the sharing of information among members of the community.

Within this framework, key *policies* must be formulated that seek to overcome residual barriers to the sharing of key data. DoD Directive 8260 constitutes an important initial step, but more needs to be done. This establishes the context for two key areas: the changes in *processes* that are needed to support key institutional activities and the *tools* that are required to facilitate the complex actions associated with the collection, conversion and sharing of information.

Finally, key *products* are needed to explore and refine our understanding of key new data concepts (e.g., the analytical baseline). The observations and recommendations that follow are organized using this construct.

Current DoD Best Practices

- Funding VV&A as part of program development
- Seeking and accepting community input in creating and implementing definitions, standards and practices
- Use of existing Commercial Off-The Shelf (COTS) and Government Off-The Shelf (GOTS) that enable current technical and analytic processes
- Migration to a "Task, Post, Process, Use (TPPU)" concept in intelligence to focus on the analyst and data from the previous "Task, Process, Exploit, Disseminate (TPED)" process focused on the collector
 - Analysis should follow same path
- Implement the DoD Discovery Metadata Standard (DDMS) as a method of highlighting the criticality of metadata in order to make data accessible to the user
- Infrastructure and services to permit users to find and retrieve data such as Net Centric Enterprise Services (NCES), Horizontal Fusion, shared spaces and metadata catalogues/registries.
- AMSAA and JDS data maintenance and data management concepts



The workshop identified a number of trends and methodologies seen as "best practices," not only for DoD analytic and study data but worthy of emulation in other arenas.

In particular the trend toward funding VV&A of models and data as part of program development rather than an after-the-fact add-on is a policy change that is contributing to improved credibility of tools and the data required to use them. A process improvement has been the government's desire to include community input in the creation of standards and new practices. This workshop is one beneficiary of that attitude. In the tool arena, the use of existing COTS and GOTS tools enable current processes and help the community in self organizing to meet important objectives in the data arena. An intelligence process improvement has been the movement from TPED processes to TPPU methods that shifts the focus from the producer to the user; this same shift in emphasis in the analytic data area is enabling the creation of new tools and processes to focus on the analyst and the decision maker, rather than the individual producer of data and will help in the long run to establish appropriate data methodologies and practices of greater benefit to the community.

The workshop reinforced the use of Metadata as a key to improvement in data management. An excellent tool in this area is the DoD's Discovery Metadata Standard that is already available to help facilitate data management and exchange. Implementation will make data more accessible to the using community. Infrastructure and services that already exist, such as Centric Enterprise Services, Horizontal Fusion, shared spaces and metadata catalogues and registries will all assist in moving the DoD data enterprise to meet the objectives set for the community. AMSAA and JDS have also developed and implemented process and product enhancements including a number of data maintenance and data management concepts that were recognized by Workshop participants as particularly strong and useful.

Candidate Best Practices (1 of 2)

- **Policies**

- Adopt incentives (and weed out disincentives) for data sharing across services, commands, and communities
- Create incentives and methods to shift the power in data management and accessibility from the data generator to the data user

- **Organizations**

- Appoint a DoD data curator
- Establish appropriate committees and working groups to develop data naming conventions, ontologies, common semantics and syntax

- **Tools**

- Develop standard algorithms to accompany standard data as part of an expanded Metadata approach
- Establish methodologies to ensure M&S data are used in a manner that is consistent with its original definition and context



In addition to current best practices, speakers and other workshop participants identified a number of other candidate best practices that could be adopted for DoD use. For example, in the policy arena, the Department could provide incentives for data sharing to encourage culture change in the Department. The benefits of having good data shared across service, command, community and activity lines far outweigh the costs to individual organizations of providing such data. The Department needs to develop a range of carrot and stick alternatives to move the community toward facilitated data sharing. Using available systems and technologies to organize around and empower the data users vice the data generators would go a long way to improve practices in this area.

With respect to organizations, a central authority and repository for all DoD data was seen as not only not possible due to the literally thousands of activities that generate, convert and use data for different purposes, but not desirable as well. Instead, the national library system is a model for what could be a more effective data enterprise. A central data curator in conjunction with a senior joint oversight or steering group could identify a unilateral structure for core and common data requirements, and assign organizational responsibility to the appropriate organizations for the maintenance of these broadly required items. Such a joint DoD data committee would be made up of representatives of the services, OSD principles and Joint staff assisted by appropriate technical expertise from the FFRDCs, academia and industry.

The curator and committee would work top level issues, such as establishing definitions, metadata standards, a controlled data vocabulary, and identifying new technologies.

In the tool area, the Department should consider developing standard algorithms and methodologies to ensure modeling and simulation data are used in a manner that is consistent with its original definition and context. Standardized language and metadata would facilitate the appropriate use of data.

Candidate Best Practices (2 of 2)

- **Processes**

- Incorporate library science techniques in repository design
- Improve visibility and accessibility of DoD data resources
- Establish formal mechanisms to simplify accreditation and adaptation of accredited data for analytic purposes
- Seek out and support opportunities to improve dialogue between analysts, model developers and operational experts
- Establish formal mechanisms to simplify accreditation and adaptation of accredited data for analytic purposes
- Encourage implementation of formal data audit trail/pedigree processes
- In most cases, standardize and organize data processes within provider organizations rather than across organizations, connected by an effective external mechanism for data sharing

- **Products**

- Develop a controlled data vocabulary (keywords) to be used throughout the analytic and data enterprises
- Provide a unilateral structure of core and common content—order of battle, METL, weapon effects, threat representations—and assign organizational responsibility for maintenance



A number of process improvements could also enhance DoD data practices. These include the incorporation of library science techniques. In addition, processes to improve visibility and accessibility of data resources should be considered along with the establishment of formal mechanisms to simplify accreditation and adaptation of accredited data for other analytic or operational purposes. Likewise, the Department should continue to sponsor activities that improve dialogue between those who develop data and those who use data for analytic or other purposes. These include activities such as this MORS workshop but could also incorporate a range of initiatives designed to connect the data production and data application communities together more effectively.

Other process best practices could include implementation of formal data audit trails for data organizations gathering, creating and using data, in order to guard against misapplications and help inform future users of the data about its character and limitations. Peer reviews, collaborative working groups, metadata standardization and other techniques could all assist in this activity.

As noted earlier, standard processes are most appropriate and effective within organizations rather than across organizations. However, the community needs to provide a forum for sharing these processes so that organizations can select from an array of processes that work to ensure internal standardization occurs. The Department can facilitate these exchanges as well as put in place useful business rules that better connect organizations and facilitate data exchange.

Best practice products include a controlled data vocabulary and a unilateral structure of common and core data as noted earlier in the organizational best practice section.

Observations (1 of 4)

- **Environment/Culture**

- The data culture requires a *fundamental* change (e.g., power must be perceived as deriving from *sharing* vice *hoarding* data)
- Must *break down barriers between the diverse communities* who are part of the data enterprise
 - Meetings, coordination efforts, collaborative study efforts, socialization
 - Initiate actions to address existing security barriers (government/industry, stealth/SAR, US/coalition)



Environment/Culture. Historically, many individuals in complex institutions have believed strongly in the mantra that “knowledge is power” (particularly if it is hoarded by those individuals). In most of the plenary presentations, a new paradigm was espoused: “power is derived from *sharing* data.” This was particularly evident in the presentation by John Osterholz which promoted a new process (“TPPU”) in an attempt to “give power to the edge.”

However, most of the working groups noted that to effect that cultural revolution it will require a basic change in the incentives and disincentives that currently characterize the data enterprise. In particular, there is a need to overcome the *fears* that impede the sharing of data (e.g., fears of misuse, misunderstanding and adverse consequences). This is a challenge that must be addressed by the senior-most levels of DoD.

In addition, most of the working groups observed that DoD would have to share data with an increasingly large and diverse set of other organizations to perform credible analyses (e.g., in support of analyses of counter terrorism, the DoD would support the Department of Homeland and the Department of State; in support of analyses of peace support operations, DoD would work closely with other government agencies, coalition partners, international organizations and NGOs). Consequently, actions should be initiated to break down the cultural barriers with these diverse communities. As one preliminary step, it would be highly desirable if MORS were to conduct one or more multi-community workshops on data sharing to support “New World Disorder” operations, at the *unclassified* level. It might be appropriate to conduct such workshops with other professional organizations that are sensitive to the cultures of the other key participants. For example, if a workshop were to be held on better data practices in the context of peace operations, it might be advisable to team up with the Cornwallis Group, which has strong ties both to MORS and the peace operations community.

Observations (2 of 4)

- **People -- Analysts**

- Need to adopt curricula and programs to *enhance education and training for the military operations analysts* that emphasize the criticality of good data practices in the analysis process

- **People – Decision Makers**

- The commitment of senior decision makers to address the data problem is helpful and should be institutionalized
- Decision makers should have access to data-related questions that they could pose to the analytic team (see next slide for strawman)



People. Only one panel, WG 2, emphasized the Education and Training (E&T) challenges that the analyst faces in dealing with the data problem. However, the Synthesis Panel perceived the E&T of providers and users of data to be a critical enabler. This includes courses to sensitize the students to: (1) the multi-disciplinary problem (e.g., how to deal with information that addresses social science factors (such as, economics, demography, sociology, anthropology, political science, as well as classical military data); (2) the data management problem (e.g., the need for discipline in generating meaningful metadata); and, (3) the technical problem (e.g., how to deal with databases that routinely exceed a terabyte). To deal with these issues, it would be appropriate for the military schools and universities to develop curricula that teach the improved data practices required to support future defense analyses.

Finally, it must be stressed that education needs to go both ways — decision makers will require education as well as the analyst. In particular, it is vital that decision makers be educated about the data problem and the level of commitment that is required to address the problem. One way of signaling that commitment is for the enlightened decision maker to pose key questions on the data issue at key moments during the life cycle of an analysis (see the following slide). That will send the message that data matters and that an analysis is not acceptable if it has not addressed the data issue systematically throughout the entire analysis process.

Data Questions to the Analyst

- **Prior to the Study**
 - What data/information (or authorization) do you need from me?
 - What do you need from the community? How are you going to obtain it?
- **At the Initial Review**
 - What do you perceive the data issues to be?
 - What plan have you generated and coordinated to address these issues?
 - What barriers (if any) do you perceive in order to access, collect, generate or convert the data required? How does your data plan address these barriers?
- **At the First Iteration**
 - Are you adhering to your data plan?
 - If not, why not?
- **At Final Report**
 - What key lessons did you learn with respect to data?
 - What steps are you taking to make your data *accessible and useful* to the greater community?



The above slide (referred to in previous slides) provides a strawman set of questions that the decision maker should pose to the analysis team at each crucial milestone of an analysis. The slide is motivated by an analogous set of questions contained in a Code of Best Practice produced by the NATO Research and Technology Organization (RTO)-sponsored research group (SAS_026).

At the outset, the decision maker must appreciate that he may hold the key to accessing needed data by providing needed authorization. Second, if the process is to be disciplined, it is vital that the decision maker require that a meaningful data plan be generated (and adhered to!) that supports the collection, generation, conversion and verification and validation of needed data. Ultimately, since the sharing of data is the key to an effective analysis enterprise, it is vital that the decision maker pose the question: "What steps are you taking to make your data accessible and useful to the greater community?"

The above list represents an initial effort to generate the questions for the decision maker to pose. It should be regarded as a strawman that should be refined in future MORS workshops or codes of best practices.

Observations (3 of 4)

- **Organization**

- Useful to adopt organizational mechanisms to encourage Interagency, International cooperation on data sharing
 - DoD Data Management Committee
 - Incentivize use of metadata, data sharing

- **Policies**

- Need to reassess existing policies which restrict the flow of data, information across institutional barriers -- *rebalance* security concerns and the “need-to-know”
 - Current “need-to-know” policy embeds a presumption of guilt, vice innocence

- **Tools**

- Must expand the analyst’s “tool chest” to support data collection, generation, conversion, V&V and visualization



Organization. The Synthesis Panel concluded that if defense analysis is to be improved through better data practices, it will require enhanced interagency and international cooperation. This requires the establishment of organizational ties with key government agencies (e.g., Department of State, DHS) and international organizations (e.g., NATO) to address the issues associated with data sharing.

Policies. Several panels observed that existing policies severely restrict the flow of data and information within DoD, between DoD and industry, and across the institutional barriers that separate the participating communities. For example, in counter-terrorism analyses, it may involve organizations at the federal, state and local level, as well as commercial entities (e.g., Chemical Transportation Emergency Center (CHEMTREC)). Current law, security policies and procedures, and concerns about proprietary information significantly restrict the dissemination and sharing of critically needed information among those participants. There is a need for the government to undertake a fresh rebalancing of security concerns and the “need-to-know.” As noted by one of the participants at the workshop, the existing “need-to-know” process is based on the presumption of guilt vice innocence.

Tools. Several of the working groups identified tools that would be of value in collecting, generating, converting and verifying and validating data (e.g., tools to support automated translating of data into the formats required by the analyst). In addition, one of the plenary speakers, Dennis Hill, HIU, emphasized the importance of developing visualization tools (e.g., VISTA) that can be used to rapidly synthesize and display the most critical data to the decision maker. Efforts should be made to refine and disseminate those tools throughout the community. In addition, those tools should be refined continually to reflect the “lessons learned” in their application to major analyses. Moreover, their efficient and effective use should be taught at the analyst courses cited earlier.

Observations (4 of 4)

- **Processes**

- Need a data support business process that exploits strengths (e.g., encourages the generation of metadata) and ameliorates weaknesses (deals with disincentives such as proprietary concerns)
- A NATO Studies, Analysis and Simulations (SAS) Panel to develop an alliance Code of Data Best Practices (analogous to C2 Assessment CoBP) should be considered

- **Products**

- Need to conduct pilot studies to shape guidance to clarify the desired attributes of the Analytical Baselines
- A significant product goal is a networked data library system that catalogues and links new and existing analytic data in archived, accessible and validated data repositories



Processes. Several participants recommended that a data support business process be developed that reinforces key strengths (e.g., encourages the generation of metadata) and seeks to ameliorate potential weaknesses (e.g., seeks to address disincentives to sharing, such as proprietary concerns). As a foundation for this effort, the sequence of events should follow the recommendation of WG 4. As they proposed, organizations should first be encouraged to get their *internal* processes in order. Once they have accomplished that objective, they are in a sound position to implement processes that support *external* sharing of data.

Members of the workshop's Synthesis Panel had participated on NATO Studies, Analysis and Simulations (SAS) panels to develop Codes of Best Practice (CoBP). These included CoBP's for C2 Assessment and Assessment of OOTW. [Note: as an illustration, selected examples of best practices for data from the C2 Assessment CoBP are provided on the following slide]. Those forums provided the time to formulate and validate best practices and included the perspectives of key NATO allies. In view of the success of those activities it is recommended that an SAS panel be convened to develop a CoBP for Improving Defense Analysis Through Better Data Practices. Alternatively, such an endeavor could be done under the aegis of the The Technical Cooperation Program (TTCP) or the American, British, Canadian, Australian (ABCA) program.

Products. As a point of departure, the Operational Availability study is being employed as a test case to clarify the desired attributes of the Analytical Baselines. It is recommended that a series of pilot studies be performed (addressing a broad range of strategic analyses) to ensure that the needs of the full spectrum of institutional assessments will be met by the evolving Analytical Baselines.

In addition, the workshop revealed that initial repositories and data warehouses are being established to support strategic assessment needs. It is recommended that these be expanded to meet the data needs of a broader range of assessment activities (e.g., simulation based acquisition).

NATO CoBP on C2 Assessment

- **“The analysis team must determine**
 - What data are needed in which structure
 - Who owns these data
 - Security issues
 - Costs to buy, collect or generate data”
- **“In the absence of needed data, it is good practice to use the knowledge of subject matter experts to generate the needed data”**
- **Metadata -- “information about information”**
 - “The source of the data, the reliability and assorted assumptions and constraints *must* be captured in standardized metadata sets”
- **Common Data Infrastructure --**
 - “Adhere to data engineering principles to contribute to data reuse”
 - “Archive data in retrievable form using standardized metadata sets”



To illustrate the level of guidance that is provided in a representative Code of Best Practices (CoBP), the above slide summarizes selected entries from the Analyst’s version of the Data section of the NATO CoBP for C2 Assessment. In that activity, a detailed, multi-hundred page CoBP was created, and supplemented with two products: a highly abbreviated 15 page version for the Analyst that was available for rapid reference, and a compressed version of the CoBP for the decision maker. If a panel is created to develop a CoBP for Improving Defense Analysis Through Better Data Practices, it would be useful to create an analogous set of products.

Summary

- **The data problem is enormous ... and exponentially increasing in size and complexity**
- **The community has taken significant *initial* steps to address the problem; e.g.,**
 - Issued new *directives, instructions* (e.g., 8260)
 - Created new *organizations* (e.g., JADMSC)
 - Formulated a *framework* based on the concept of enterprise, community of interest and private data
 - Promulgated new *tools, standards* (e.g., DDMS)
- **However, further substantive improvements requires additional action:**
 - *Changing the culture* by implementing incentives and overcoming disincentives
 - *Educating and training* users and providers of data — as well as decision makers
 - Implementing new *processes* with a focus on metadata



There was broad agreement among all the participants at the workshop that the data problem is enormous ... and increasing in both size and complexity!

The workshop served to showcase several significant initial steps that the community has taken to address the problem. First, there was broad agreement that the new directives (DoDD 8260.1) and instructions (DoDI 8260.2) provide a useful foundation for attacking an important segment of the problem. There was anticipation that the proposed Analytical Baselines will serve to enhance the speed and quality of future strategic analyses. In addition, there is hope that the Joint Analytic Data Management Steering Committee (JADMSC) will provide an effective forum for overseeing the creation and evolution of these Analytical Baselines. Further, the framework promulgated by OASD(C3I) to conceptualize the problem (i.e., enterprise, community of interest, private data) appears useful along with the new tools and standards that they are developing (e.g., the DoD Discovery Metadata Standard (DDMS)).

However, in order to make additional substantive improvements, it is vital that we take several challenging steps. The most important of these steps is an initiative to transform the culture of data from one of hoarding to one of sharing. To do so, steps must be taken to dispel the fears that permeate the community (e.g., fears of misuse, misunderstanding and adverse consequences). This initiative must be undertaken and sustained at the highest levels of leadership. Second, one must address the “people” issues. This entails educating and training the users and providers of data — and the decision maker! Finally, if there is one key technical issue to be worked it is that of metadata. This theme was mentioned by most of the plenary speakers and it constituted the baseline recommendation of the working group on data management. It was observed that the problem is hard and that we must begin to address it seriously immediately.

Scorecard

- **Accomplishments:**

- The Workshop made progress against four key objectives:
 - Review current data processes (****)
 - Identify best practices and standards (***)
 - Identify and assess potential sources of data (***)
 - Educate the community on the new data directive (****)
- To paraphrase Philip Roth: “Now, we are ready to begin...”

- **Residual Challenges**

- Need to reach out to a broader community (e.g., inter-agency, coalition) to address the data problem
- Need to generate a Plan of Action and Milestones (POA&M) to focus, mobilize the community
- Regard this workshop as only the first in a sequence



A scorecard to summarize the accomplishments of the event and the residual challenges that confront the community helps put the workshop in perspective.

Accomplishments. The workshop sought to achieve four objectives: 1) review current data processes; 2) identify best practices and standards; 3) identify and assess potential sources of data; and, 4) educate the community on the new data directive. In the view of the Synthesis Panel, the workshop made excellent progress (“four stars”) on two of those objectives (i.e., review current data processes and educate the community on the new data directive), and good progress (“three stars”) against the remaining two objectives (i.e., identify best practices and standards and identify and assess potential sources of data). Given the highly complex, multi-dimensional problem of interest, these are major accomplishments.

However, it must be understood that this constitutes a beginning, not an end. Our status is reminiscent of Philip Roth’s Portnoy at the conclusion of his book “Portnoy’s Complaint.” After a difficult, soul-wrenching confessional, his psychiatrist concludes the book with the words “Now, we are ready to begin...”

Residual Challenges. Since we are now “ready to begin,” we believe that there are three major initiatives that the community should undertake. First, there is a need to reach out to a broader community to address the data problem. This includes enlisting the aid of the inter-agency and alliance/coalition communities. Second, there is a need to generate a POA&M to focus and mobilize the analysis-data community. Such a product will help coordinate diverse actions and provide the discipline needed to ensure steady progress. Finally, MORS must recognize that the complexity and size of the problem is such that a single workshop is inadequate. Plans should be initiated to conduct follow on workshops to address many of the critical issues identified during this meeting.

BACKUP SLIDES



Key Issues (noted by JADM IPT)

- No single solution; requires comprehensive change to culture, people, organization, policies and tools
- DoD intelligence is migrating from a collector centric "Task, Process, Exploit, Disseminate (TPED)" concept to a "Task, Post, Process, Use (TPPU)" process focused on the analyst and data user
 - Analysis should follow same path
- Requires *enhanced education and training for military operations analysts* as well as continued commitment by senior decision makers to address the data problem
- A key organizational step would be to adopt library science solutions to include appointing a data curator for DoD
- Data policies should be aimed at creating incentives (and weeding out disincentives) for data sharing across services, commands, and communities
 - Means evolving approaches to shift the power in data management and accessibility from the data generator to the data user
- Metadata will be a key tool—need to develop standard algorithms to accompany standard data as part of an expanded Metadata approach



The workshop identified a number of trends and methodologies seen as "best practices," not only for DoD analytic and study data but worthy of emulation in other arenas. The Joint Analytic Data Management Integrated Process Team (JADM IPT) selected a number of these for presentation to the JADM Steering Group.

In particular the trend toward funding VV&A of models and data as part of program development rather than an after-the-fact add-on is a policy change that is contributing to improved credibility of tools and the data required to use them. A process improvement has been the government's desire to include community input in the creation of standards and new practices. This workshop is one beneficiary of that attitude. In the tool arena, the use of existing COTS and GOTS tools enable current processes and help the community in self organizing to meet important objectives in the data arena. An intelligence process improvement has been the movement from TPED processes to TPPU methods shifts the focus from the producer to the user; this same shift in emphasis in the analytic data area is enabling the creation of new tools and processes to focus on the analyst and the decision maker, rather than the individual producer of data and will help in the long run to establish appropriate data methodologies and practices of greater benefit to the community.

The workshop reinforced the use of Metadata as a key to improvement in data management. An excellent tool in this area is the DoD's Discovery Metadata Standard that is already available to help facilitate data management and exchange. Implementation will make data more accessible to the using community. Infrastructure and services that already exist, such as Centric Enterprise Services, Horizontal Fusion, shared spaces and metadata catalogues and registries will all assist in moving the DoD data enterprise to meet the objectives set for the community. AMSAA and JDS have also developed and implemented process and product enhancements including a number of data maintenance and data management concepts that were recognized by Workshop participants as particularly strong and useful.

Potential Data Sharing Incentives

- **System changes to move focus from data generator to data user**
 - Identify and support processes, approaches and technologies that enhance user role
- **Security Issues**
 - Need to know rules (TFEL example: DARPA, CENTCOM and IDA)
 - Stealth SAR regulations
 - Use of surrogates are punishable by fines and prison (if stealth addressed, must be done in SAR environment or not at all)
 - JADM or selected service and joint offices could certify broadly-based "need to know" for mission and campaign analysis; work with SAF/AQ to resolve stealth M&S
- **Economic incentives**
 - Good data has economic value; Department could facilitate market
 - Designate organizations responsible for generating specific types of data for the broader community and compensate accordingly
 - Industry has indicated interest in paying for basic scenario and force data
 - Would need internal regulation to ensure appropriate applications and address abuses
- **Leadership support**
 - Military culture aligns organizational effort to support lawful initiatives of assigned leadership; leaders must clearly support "best warfighting capability" over internal goal of increasing organizational size and resources
 - Awards and recognition program
 - Promote individuals based on community vice organizational contributions



The JADM Steering Group also asked the workshop leadership to suggest incentives that the Department and the Services could consider to help the community move toward better Data Practices. The list on this slide was based on workshop discussions and summarizes some of the ideas identified by the working groups in general or by individual participants. They include changing system focus to the data user, fixing security barriers, providing economic incentives for these practices and helping to generate leadership support for data sharing vice data hoarding.

Data Process Definitions

- **Data Conversion** - The translation of data from one representation format to another including modifications of semantics (meaning) and/or syntax (storage representation). Includes:
 - **Format** - Changing the "bits and bytes" arrangement of data for specific application use.
 - **Mathematical** - Translation of data using a function which can be expressed as an algorithm. This embodies the mathematical definition of "function" where a set of inputs give one unambiguous output
 - **Mapping** - Translation from one set of discrete set of data values to another set of values by either objective rules or subjective judgments
 - **Aggregation** - Translation from a granular set of data values to a more abstract set of values
- **Data Creation** - The process of deriving and storing unique attributes of defined variables for a specific purpose within an enterprise
- **Data Management** - The set of processes and standards used to define, collect, store, share, apply and direct the use of data within an enterprise
- **Data Security** - Providing the appropriate level of protection to data and information to protect from unauthorized access and unauthorized changes. Accomplished by DoD classification levels, access control or other protective labeling and controls (Privacy Act, Proprietary). Includes systems, procedures and training
- **Data Sharing** - The process of providing existing analytic data among the levels of an organization, different agencies or services. Issues/problems/barriers include proprietary nature data, bureaucracy, lack of confidence, timeliness, regulations/laws, fear of inappropriate use, security or lack of security
- **Data Surrogation** - Use of fabricated data to representative data not otherwise available for use in studies, exercises and experiments (e.g., estimating performance data for systems that have not yet been invented, by using a multiple of a current system's attributes and performance)



Data Type Definitions

- **Common Data** - Parametric representation of common characteristics or generally accepted standards
- **Community of Interest Data** - Shared among organizations pursuing similar goals
- **Core Data** - Essential elements of data needed for virtually all applications usually without conversion (equipment and force descriptions, physical properties, etc.)
- **Custom Data** - Specifically tailored to a scenario and not of sufficient scope for general use
- **Enterprise Data** - Information representing an entire organization
- **Private Data** - Internally developed for sole use by the organization
- **Scenario** - An account or synopsis of a projected course of action or events. ... the focus of scenarios is on strategic and operational levels of warfare. Scenarios include information such as politic-military contexts and/or backgrounds, assumptions, operational objectives (threat and friendly), major force arrivals and planning considerations. (DoD Directive, Dec 6, 02)
 - **scenario**: Set of data and information needed to assess the capability of a force or force element to accomplish an objective.
- **Threat Data** - Parametric representation of potentially hostile systems



Data Quality Definitions

- **Accessibility** - Ease of obtaining the data by the using organization
- **Accreditable** - Ability to be approved for a specific application by an appropriate source or authority
- **Affordability** - Utility of acquiring the data for use
- **Availability** - Breadth of community knowledge as to data location and pedigree
- **Credibility** - Extent to which data is believable based on its source or pedigree
- **Data Purity** - Measure of a data set's accuracy and credibility



Data Quality Grading Standards

- **Data Shelf Life** - The period of time over which the data remains valid/accurate
- **Holistic Data** - Data representing the whole rather than components or processes
- **Institutionalized Data** - Data utilized within an enterprise and available to all elements; normally gains legitimacy due to wide-spread acceptance and use
- **Interoperable** - Useable by and credible to various applications
- **Suitability** - Applicability of the data to the specific uses desired
- **Validatable** - Data declared to be useable by predictive methodologies
- **Verifiable** - Data declared to be accurate and representative of the original system
- **Visibility** - Ability to determine source, pedigree, and original form of data and its application



Other Definitions

- **Analytic Agenda** - Current plan for providing timely analytic results to DoD leadership questions
- **Asymmetric Warfare** - Conflict characterized by non-linear or multi-dimensional warfare
- **Authoritative Data Source** - Recognized source of credible data for specific applications
- **Canonical Standards** - A set of standards which supersede all others and apply broadly to all organizations
- **Effects Based Operations** - Military plans which focus on the end-state rather than the specific functions of execution
- **Metadata** - Data describing other data
- **Metamodel** - A model that describes a model -- the class variables of a particular kind of model or a way of discovering a model. In the data enterprise, metamodels help define data scope and attributes necessary for applying a model and obtaining useful results.
- **Non-Linear Modeling** - Models or representations of military functions or applications that attempt to capture system non-linearities
- **Ontological Development** - Relating to real-world or existing characteristics
- **Repository** - A central location providing access to data archives
- **VV&A** - Verification, Validation and Accreditation



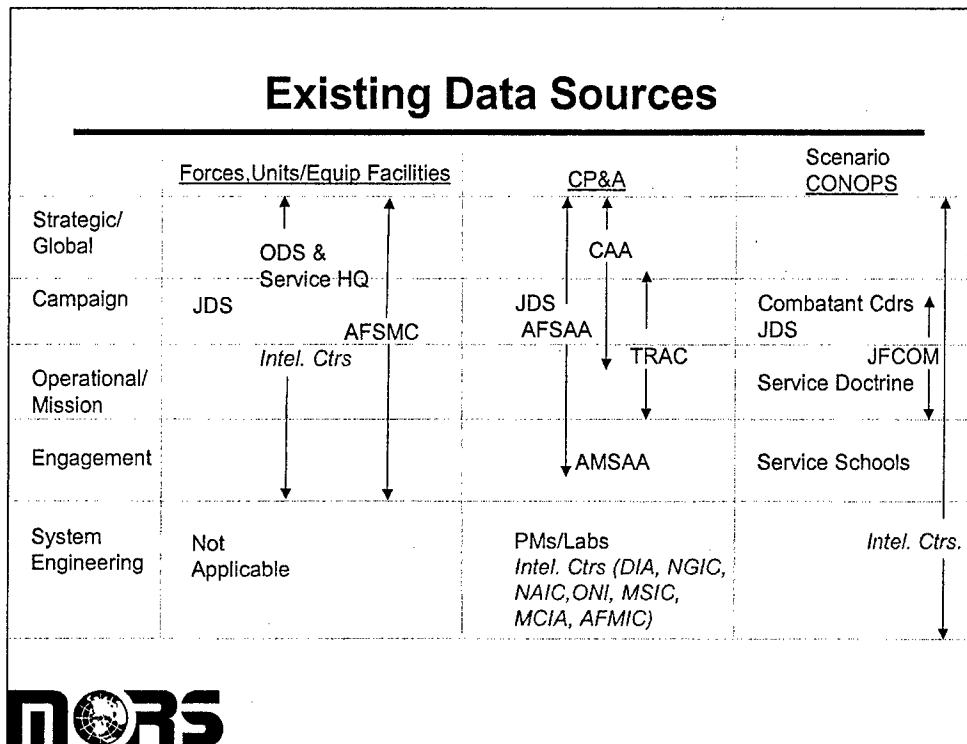
KEY:

- Analysis
- System Data
- Data Methodology
- Tactics CONOPS

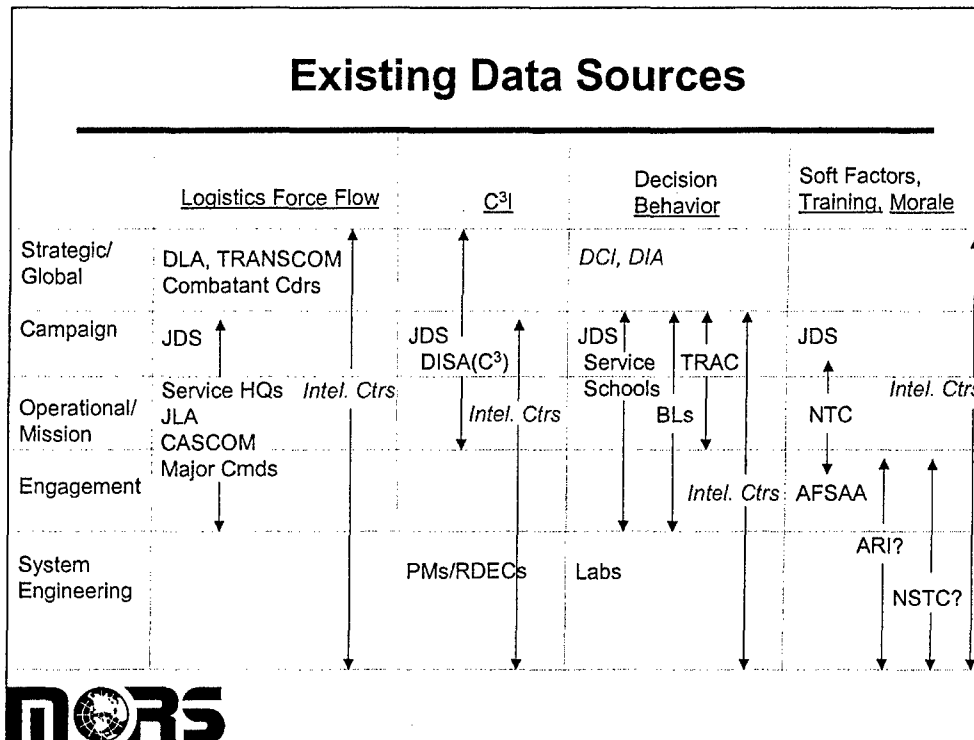
Just for one organization!



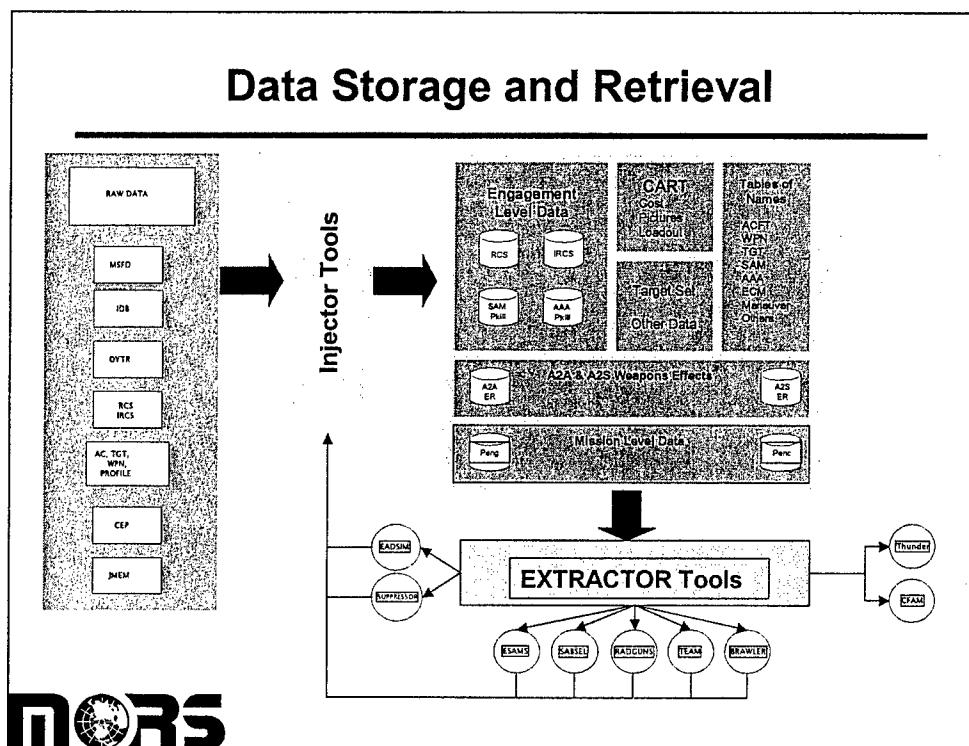
46



This chart and the next examine levels of warfare analysis, and identify some of the organizations that provide data in the major categories relevant to M&S. Note, intelligence agencies that provide data on green, red and brown systems are depicted in red (and are italicized).



This slide is an extension of the last in other areas of interest.



This is an example of the type of system analytic and study organizations are attempting to develop to facilitate their activities. Different organizations have achieved different levels of success in standardizing their internal processes, but these blocks and arrows provide context to the normal types of operations conducted by each one of these organizations. In a sense, JDS tries to provide a similar system at the macro level within the Department of Defense.

Technologies Discussed/Briefed

- **Semantic web technologies**
 - Enhance discovery process
 - Link ontologies
 - Enable semantic joins
 - Allow machines to read and inference about web content
 - Represent data in XML/RDF/OWL
 - ISSUE: Scalability for inferencing
 - ISSUE: Ontology registry required



The Technology Working Group reviewed a number of interesting and helpful technologies. This graphic and the remaining slides provide thumbnail descriptions of the purpose or contributions of specific technologies or applications. In addition, the slides include issues developed by the working group that require understanding and that should be addressed by organizations seeking to adopt these technologies to their data support requirements.

Technologies Discussed/Briefed

- **Modular engineering models**
 - Streamline data creation
 - To rapidly populate databases
 - Empower data end user
 - Respond to changing user needs
 - Capture/output of anything calculated
 - Use model to generate specific data set automatically
 - Embed in other simulations
 - Reusable
 - Modular
 - Hyperlink to engineering descriptions
 - Standardize on MatLab/Simulink
 - ISSUE: Configuration management/data integrity



Technologies Discussed/Briefed

- **Distributed application integration**
 - Perform data normalization
 - Perform data validation
 - Provide capability to enrich data
 - Emerging technologies:
 - Improve visual design tools
 - Increase metadata transparency
 - Standardize XML data definitions
 - Implement more efficient web services
 - Automate access to web sites to extract & insert data
 - Perform graphic recognition
 - ISSUE: History of data use/reuse
 - ISSUE: Track of data generation/transformation



Technologies Discussed/Briefed

- **Common Warehouse**

- Metamodel**

- Continue improvement and implementation of standards
 - UML
 - XMI
 - MOF
 - CWM
 - Support
 - Stewardship/ownership
 - Browse and update
 - Security
 - Flexible categorization/annotation on field by field basis
 - Search/reuse
 - Collaborative working
 - Notification
 - Metrics both on usage and quality
 - Versioning
 - Federation of catalogs
 - Track transformation of data
 - ISSUE: Tracking data provenance
 - ISSUE: Provide a more efficient search engine for Metadata
 - ISSUE: Advanced automated feature selection
 - ISSUE: Ground truth, temporal tagging
 - ISSUE: Data reduction



Technologies Discussed/Briefed

- **Unstructured content to meaningful XML**
 - Implement portals
 - Use tools for creating XML for content as produced
 - Apply methods for automating XML for legacy documents
 - Learning based extraction
 - Rules based extraction
 - Pattern recognition
 - ISSUE: Natural language processing
 - Semantic processing
 - Computational linguistics



Technologies Discussed/Briefed

- **WebTAS**

- Visualized relational, flat, XML
- Hyperlink records
- Model/time event detection
- Solutions send out via web pages
- XML import/export



Improving Defense Analysis
Through Better Data Practices
March 25 - 27, 2003, Alexandria, VA

**MORS Workshop Outbrief Report
Working Group #1: Data Management**



**Chair: Simone M. Youngblood
Co-Chair: Jim Stevens**

Agenda

- **Membership**
- **WG #1 focus and “rules of engagement”**
- **Key issues**
 - State of the practice
 - Lessons learned
 - Recommendations
- **Substantiating information**
- **What we didn’t get to**
- **Conclusions**
- **Follow-on activities**



After introductions and review of the workshop approach, the working group addressed three key issues impacting data management: 1) the current state of practice in data management for DoD study and analysis support; 2) lessons learned from various activities in the area of data management; and, 3) recommendations from these for action by relevant organizations within the DoD. To the extent possible, the working group attempted to reference substantiating information to clarify points and recommendations. This brief will also quickly note what we wanted to cover but didn’t get to as well as conclusions and recommendations for follow-on activities.

Members

<u>NAME</u>	<u>ORGANIZATION</u>
William F. Diehl	Alion Science & Technology Corp
Virginia T. Dobey	VTC Dobey Associates, Inc
Nancy H. Dunn	Army Evaluation Center
Norman D. Edwards	DFI International
Johnny Garcia	JFCOM J9
Gray M. Gildner	Alion Science & Technology Corp
James W. Hollenbach	Simulation Strategies, Inc.
Terence Hooper	DSTL (MODUK)
Michael K. Hopkins	Computing Technologies
Stephen M. Hunt	OSD (PA&E/JDS)
Richard Eric Johnson	Unisys (OSD PA&E/JDS)
Thomas Howard Johnson	Alion Science & Technology Corp
David E. Kendrick	Veridian Engineering
Thomas A. Mahalek	US General Accounting Office
Gary L. Misch	Alion Science & Technology Corp
Robert M. Senko	Titan Corp (DMSO)
LT Phil V. Ventura	Assessments Division (OPNAV N81)
Joel E. Williamson	DFI International
Simone M. Youngblood	DMSO (JHU/APL)



Workshop members included a broad range of analysts and data managers from throughout the Services and DoD departments, as well as from industry.

WG Focus and “Rules of Engagement”

- **Data Management Working Group**
 - Objective
 - To provide insights and recommendations that will assist the DoD in managing risk associated with data.
 - Key Issue - A framework is required to enable:
 - Data discovery
 - Data access
 - Understanding data for use and reuse
- **Approach/Methodology used by the group**
 - Key presentations (MSRR, JSF, JDS, UOB, ADS)
 - Used to frame discussion, identify insights, obstacles and potential solutions



The objective of the working group was to provide insights and recommendations to assist the Department of Defense in managing the risk associated with data. A primary proposition is that any organization dealing with data needs a framework to enable data discovery, data access and the appropriate use and reuse of data.

The group reviewed some exemplary presentations on data methodologies by current efforts within the department, such as the Modeling and Simulation Resource Repository (MSRR) process established by the Defense Modeling and Simulation Office, the data approach used by the Joint Strike Fighter (JSF) development program, the Joint Data System (JDS) management methodology and others. These were used to frame discussions and focus on issues.

Data Discovery

- **State of the practice**
 - Predominant method of data discovery is through personal contacts
 - Limited and/or ineffective resource repositories
 - Limited awareness of available resources
 - Reluctance of providers to populate resource repositories
 - No clear incentives and some disincentives
 - Unfunded workload
 - Negative impacts resulting from inappropriate reuse of data
- **Lessons learned**
 - From the M&S community: A DoD wide repository effort was marginally effective due to a lack of awareness, funding and bottom-up support
 - “If you build it they might not come”



In the area of data discovery, the group determined that the predominant approach to finding relevant data to support joint studies and analysis was the use of personal contacts. While there are data repositories for a wide variety of data, the group noted that many analysts are either not aware of them or, even if they know about them, are not sure how to quickly and successfully access them.

At the same time, many raw data developers are reluctant to pass their data to the known repositories for a variety of reasons. Issues include the lack of incentive to share data and, in some cases, real barriers to the passing of information. One disincentive is the potential for such sharing arrangements to lead to unfunded workload requirements. Another is the negative impact of inappropriate reuse of data provided by the repository that puts both the organization that originally provided the data and the system or capability whose data was inappropriately applied in an incorrect spotlight.

A lesson from the current effort to provide a DoD-wide data repository for DoD models and simulations is that that effort is only marginally effective. Despite several years of top level policy and activity it continues to face challenges due to a lack of awareness, funding and bottom-up support. As one participant put it, “if you build it, they might not come.”

Data Discovery

- **Recommendations**

- Increase visibility of available data resources
 - Focus on metadata and not the data products
 - Leverage current and emerging policy focused on improving and cataloguing metadata
- Create incentives to data providers to post their metadata.
Possible approaches:
 - Write posting/metadata requirements into contracts
 - Implement commercial based business model (fee for service, etc.)
- Employ ADS (or similar) data quality grading standard and library science to discover suitable data sets
 - Peer review



One way to improve performance in this area is to increase the visibility of the available data resources. To help guard against inappropriate use as well as facilitate data sharing, the working group recommended that the repositories and associated departmental data policies should focus on metadata and not the data products themselves. Current and emerging policy aimed at improving and cataloguing metadata could have a major positive impact on the state of data management, not only for analysis, but for other data areas as well.

In particular, the DoD needs to create incentives for data providers to post their metadata in ways that facilitate community access and review. Ways to make this happen could include the addition of specific posting and metadata requirements in contracts and the implementation of commercial-based business model practices that include fee for service arrangements for data providers.

In addition, defense organizations could employ Authoritative Data Sources (ADS) or similar data quality grading standards as well as leverage library science to discover suitable data sets, metadata standards and associated policies. In particular, by posting and sharing standards, peer review would include self-standardizing activities to ensure the metadata was understood and used appropriately throughout the community.

Data Access

- **State of the practice:**

- Reluctance of providers to share data
 - No incentives
 - May be disincentives! Perceived risk to data producer
 - Releasability issues (security, proprietary)

- **Lessons learned:**

- Social, political and cultural rather than technical issues create the biggest challenges



Data access challenges were similar to those noted in the data discovery arena. Access is characterized by the reluctance of providers to share data. Lack of incentives and perceived risk disincentives make the current situation difficult, particularly for studies requiring information from different organizations and services. In addition, releasability issues associated with security and proprietary data make the study of many issues very difficult. Stealth information, for example, is an important aspect of force performance on the battlefield. Accounting for stealth in realistic ways even when the point of the study is on some other aspect of the envisioned combat activity has been complicated by current security regulations which make the use of surrogate data a criminal offense while real data can only be used in a special access environment.

The main lessons uncovered by the working group is that the social, political and cultural issues far outweigh the technical challenges to access.

Data Access

Recommendation:

- **Discover means to incentivize data providers to share their data. Possible approaches:**
 - Implement commercial based business model (fee for service, etc.)
 - Include data/data models/metadata as contract deliverables
 - Explore data interchange methodologies



Finding new incentives and applying them to the problem is the key to breaking down data access barriers. The working group suggested a number of possibilities to include the implementation of commercial based business models, such as fee for service, to help overcome the barriers. In addition, including data, data models and, most especially, metadata as part of defense contract deliverables would dramatically improve performance in this area. Finally, the group recommended the DoD explore data interchange methodologies to ensure technological barriers were eliminated.

Understanding Data for Use and Reuse

- **State of practice:**

- Poor or no metadata
 - No common metadata syntax
 - Low to no priority for provided metadata
- No uniform application of data quality grading
- Limited practice and support for reuse of data

- **Lessons learned:**

- JSF and JDS briefings highlight the necessity of understanding data within the context of the application and the critical role that metadata plays in achieving that understanding
- Air Force Pedigree Database shows resource value for process of tagging and reuse procedures



In the final discussion area, understanding data for use and reuse, the working group found the major issue to be metadata, or more appropriately, the lack of metadata and accompanying standards. For many common data sets applied across the DoD, no metadata exists. Where metadata has been developed, there is no community standard or common metadata syntax. In either case, there has been no priority placed on providing metadata to accompany database development.

Other issues include the lack of priority placed on the grading of data quality (in the case of most studies, the single quality criteria has been “best available;” unfortunately, the quality associated with “best available” has varied widely across organizations depending on their access to data sources and their internal ability to address quality in objective terms). Likewise, while data is often reused, the standards and approaches for ensuring the reuse is appropriate has normally depended on the skill of the employing analytic team.

The briefings provided to the working group highlighted the necessity of understanding data within the context of the application and the critical role that metadata plays in attaining that understanding. The Air Force has attempted to develop some internal “rules of the road” for the metadata process by invoking specific data tagging and reuse procedures. While involving significant effort, the process highlights the value of such activity in the application of data to a range of study uses.

Understanding Data for Use and Reuse

- **Recommendations:**

- Extend DoD policy requiring the creation and local retention of metadata tied to instance data
 - Organization adds only *its* metadata with reach back before forwarding to the next organization
 - Data should not be considered complete without the metadata
- Write metadata requirement into contracts
- Adopt ADS (or similar) data quality grading standard and library science to discover suitable data sets



The working group recommended that DoD policy in this area be extended to include a requirement for creation and, at least at the originating level, retention of metadata tied to instance data. In this way, each study organization would be responsible for adding only its specific metadata with reach back to originating organizations for the basic metadata. Such an approach would provide a clear descriptive trail for data and its alteration in context for specific purposes without one organization having to bear the full metadata development burden. A major feature of such an approach would be the community-wide understanding that a database would not be complete without the associated metadata.

This requirement should also be incorporated in all study and data-development contracts to ensure metadata is created, maintained and available for appropriate reuse throughout the DoD.

Substantiating Information

- **Lessons learned from the Plenary Session:**

- “If searchable description/quality info is not good, the data might as well be lost” Mr. Kurt Schnebele (NOAA)
- “All data is contextual” Mr. Vern Bettencourt, FS (HQDA)
- “Need to know where the data has been” Dr. Gary Strong (NSF)
- Dr. Richard Morris (NIH) discussed need for “Data Curator” - data repositories don’t manage themselves



The working group used a number of points from the plenary presentations to reinforce their lessons and recommendations. Mr. Kurt Schnebele of NOAA, in addressing the problems associated with weather data, stated that, “If searchable description/quality information is not good, the data might as well be lost.” He, like many other speakers, underlined the importance of data pedigree and the requirement for metadata to ensure data description and quality is clear to data users. The Army’s Mr. Vern Bettencourt, FS noted that, “All data is contextual.” Without understanding the context in which the data was generated or the use for which it was intended, a different analytic team could easily and inadvertently misuse the data. Dr. Gary Strong, a data expert from the National Science Foundation, made the same point when he stated, “You need to know where the data has been.” Without metadata, that trail is lost. Finally, Dr. Richard Morris of the National Institutes of Health pointed out the need for an effective data curator. Data repositories don’t manage themselves and someone needs to work within the community to develop and enforce standards, particularly with respect to metadata.

Substantiating Information

- **Recommendations from the Plenary Session:**

- Incorporate library science techniques in repository design
- Develop a controlled vocabulary (keywords)
- Identify a data “curator”
- Move from task/process/exploit/disseminate to task/post/process/use, Mr. John Osterholz (ASDC3I)



A number of the recommendations offered during the plenary session apply directly to data management. Certainly the analytic community should leverage the knowledge of library science in designing data repositories. This would include the development and application of a controlled vocabulary to include keywords to facilitate data identification and guide reuse. In addition, a data curator function would go a long way toward assisting organizations that generate and hold data to establish standards and functionality that facilitates appropriate data use and effective data exchange. Mr. John Osterholz, from what was then the Office of the Assistant Secretary of Defense for Command, Control, Communications and Intelligence, recommended the analytic community follow the lead of the intelligence community. In that community the effort is being made to shift control from the data supplier to the data user through the application of a “Task, Post, Process, Use (TPPU)” methodology. Data is developed and made available to the community for user application resulting in improved data exchange and much quicker sharing of relevant information. With the appropriate metadata, such a method could serve the analytic community equally well.

What We Didn't Get To

- **What important questions did you not address due to lack of time, lack of scope or expectation that they would be addressed by other groups?**
 - Tools
 - Data Management Technical Framework (roles, responsibilities and interactions)



Despite enthusiasm and many hours of spirited discussion, the working group did not cover the full menu of topics listed at the start of the workshop. Additional attention in other forums will need to be devoted to the development and sharing of common tools in order to help standardize and streamline data management processes as well as to the appropriate construction of a data management technical framework. This later discussion will need to address the roles of all members of the analytic data community to include the concept of a DoD data curator, as well as to responsibilities and interactions expected of each of the other community members.

Conclusions

- **Improving defense analysis through better data practices requires the DoD to:**
 - Create a common data technical framework
 - Improve visibility of DoD data resources
 - Incentivize the community

Metadata! Metadata! Metadata!



The working group came away from their discussions with three main focus areas. While the specific definition and detailed rules for roles, responsibilities and interactions were not generated, the group agreed that the DoD does require the creation of a common data technical framework so that when the cultural issues are resolved, the technical problems will have already been overcome. At the same time the DoD needs to take steps to improve the visibility of their existing repositories and other data resources and incentivize the community to improve sharing arrangements and focus on the development and use of metadata throughout the analytic enterprise. The key to all of this is effective implementation and use of metadata for the collection, storage and management of data.

Follow-on Activities

- **Recommended follow-on activities**
 - Barely scratched the surface on issues related to data management. Evolve a series of follow-on workshops and/or a standing technical body to address issues raised.
- **Other areas of potential focus**
 - Management of cost analysis data



This workshop barely scratched the surface of a number of important issues. While working group members were able to chart a course forward, all believe that a series of follow-on workshops and/or the establishment of a standing body within the DoD analytic community to address issues in this area would be positive steps forward. Having started the effort by focusing on the DoD analytic studies arena and addressing primarily military element descriptors and measures of effectiveness, the group believes an important next step would be to address the development and management of cost data, another area that would benefit from shared standards and common methodologies within the DoD.

Improving Defense Analysis
Through Better Data Practices
March 25 - 27, 2003, Alexandria, VA

**MORS Workshop Outbrief Report
Working Group #2: Obtaining Data from
Outside Sources**



**Chair: Jack Sheehan
Co-Chair: Phillipe Loustaunau**

Working Group 2 addressed the difficult area of obtaining data from outside sources.

Members

- William Ashman (Army ECBC)
- Ambrose Brennan (PA&E JDS)
- Michelle Busbee (NAIC)
- John Christakos (OSEC)
- Vic DiRienzo (Alion)
- David Drake (NAIC)
- Maj Hoot Gibson
- Bernard Harvey (SAIC)
- Steve Hix (PTI)
- Ollen Landrum (NAIC)
- LtCol Eric Holdaway (J-8 PEAD)
- Frank Holland (PA&E JDS)
- Philippe Loustaunau (DFI International)
- Charles Metcalf (HQ Dept. Army)
- Raymond Miller (AFSAA/SAA)
- Jack Sheehan (DOT&E/DMSO)
- Gene Visco, FS (Visco Consulting)
- Dan Kim (DFI International)
- Bruce Esken (SAIC)
- Phil Barry (Mitre)



Members were led by Jack Sheehan and Philippe Loustaunau and included an excellent cross section of the DoD, industry experts and analytic practitioners.

Obtaining Data from Outside Sources

- **Issues, incentives with outside data**
 - Visibility, availability, suitability, access
 - Credibility, timeliness, affordability
 - Planning, producing, paying
- **Scope of discussion**
 - Requirement: Scope and framework
 - Content: Core (enterprise), common (COI), custom (private)
 - Policy: Business model and budgeting
 - Technology: Capabilities and constraints
- **Approach/Methodology used by the group**
 - Two case studies (TMAP, JMEM)
 - Four topics discussed in three sub-groups with cross-briefing



The group divided issues and incentives associated with outside data into three general areas: the first included visibility, availability, suitability and access to the outside data. The second addressed quality in terms of credibility, timeliness and affordability of the data. The third area had to do with planning, producing or paying for the data.

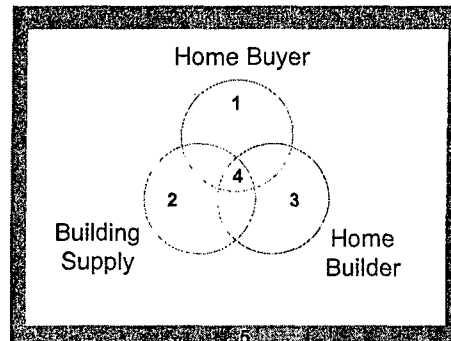
Discussion of these issues ranged from the requirement for the data in terms of scope and framework; the content of the data with respect to its commonality with the rest of the community; the policies governing how the data is obtained; and finally, the technological capabilities and constraints associated with obtaining the data.

Over the course of two days, the working group reviewed two case studies and then addressed the four discussion topics with separate groups in each of the three general areas, sharing insights through cross-briefings.

Housing Construction Metaphor

Five Data Working Groups

- 1) Data Management
- 2) Obtaining External Data
- 3) Creating Data Internally
- 4) Transforming and Employing
- 5) Data Technology



To set the stage for the discussion, the group used the housing metaphor. Just as a home buyer turns to a home builder to create the house of his or her dreams and the home builder turns to various locations to obtain supplies, the DoD study customer turns to the analytic community for studies and the analytic community in turn seeks out data suppliers as a basis for building relevant studies. The question is where is the Home Depot for data and what data should be provided on a common basis to all study activities and what data is unique enough that it requires tailoring for the specific study question at hand? These questions will be dealt with in more detail on a later slide.

Purpose and Context: Two Broad Circumstances

- **Normal Operations**
 - Decisions: JSCAP, TOA, force structure, acquisition milestones
 - Tempo: Deliberate planning
 - Paradigm: Invest in capability
 - Fame: Go slow, finish first
 - Shame: Haste makes waste
 - Objective: Get it right with all deliberate speed
- **Examples**
 - Combatant Commander deliberate planning
 - Service Chief personnel, training, spares planning
 - Joint Chiefs out-year force mix planning
 - Acquisition Chiefs new airframe, hull, chassis AoA



The group determined that there were two broad study contexts that could divide the analytic communities' normal effort. The first is the normal operations type studies that address broad questions in a deliberate way. These include the combatant commander's deliberate planning process, the services annual planning for personnel, training and spares and even the acquisition community's major studies for selecting next generation equipment.

Purpose and Context: Two Broad Circumstances

- **Response Operations**

- Decisions: Crisis management, force selection, opportunity branches and sequels
- Tempo: Crisis action planning
- Paradigm: Buy services
- Fame: Speed kills
- Shame: Delay is death

- **Examples**

- Consequence management: Air Traffic Shutdown on 9/11
- Force selection: For Operations Enduring and Iraqi Freedom
- Tech opportunity: Tactical internet, wireless communications



The second broad type of study activity is the quick response analyses, which are crisis driven or are questions that need fast resolution. These include real time operations and force selections as well as decisions to take advantage of break-through technologies.

Sequence of Success: Five Key Elements

(1) Requirements Engineering with Users

- The more complete the better

(2) Social Engineering with Stakeholders

- Must enlist voluntary cooperation

(3) Technical Engineering

- System/technical architecture should be Technology Readiness Level 7

(4) Information Engineering

- Demand pull (users pull products, products pull tools)

(5) Inventory Management

- Stock components that can be tailored for custom assembly



Five steps in either type of analysis are harbingers of success. The analysts need to conduct requirements engineering with the users to make sure they understand what the customer really needs. Social engineering with all the stakeholders is also required, since they are generally the ones who provide the data or must come up with the resources to implement the solutions. Technical Engineering is the third step, since for a program or solution to be a success, it must be technically feasible. For many DoD studies, this means a system must be at the Technical Readiness Level 7 to ensure a near-term fielding option. The next step is the information engineering, using the customer to pull the products and the products to determine the tools to be used. Once this is accomplished, the analyst needs to conduct solid inventory management, maintaining data and tools that can be easily tailored and used to product the desired end product.

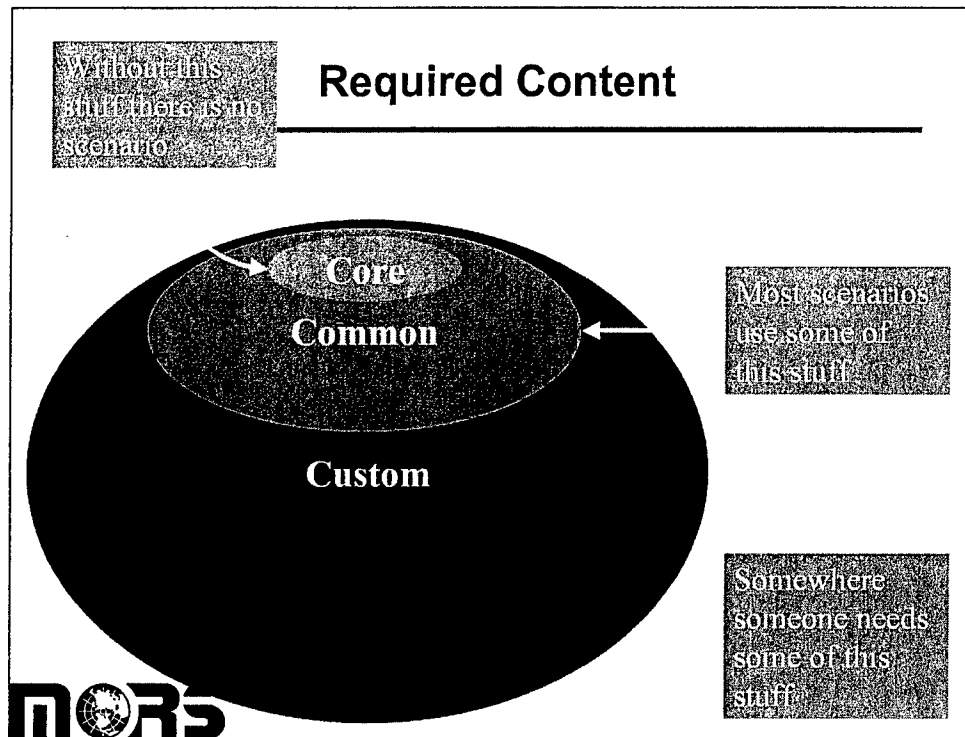
Content Issues

- **Core**
 - Information elements that capture Purpose, Identity, Context -- defines COI's
- **It's the mission that counts**
 - Defined by Mission at Level of War which in turn defines and is composed of subordinate Missions at or below the same Level of War
- **Prefabricate selected core and common content**
 - Based on difficulty, importance, frequency, mutability



Returning to the construction analogy, there are three classes of data that must be accessed in any major DoD study. These include Core, Common and Custom elements. Core data is that information that captures the purpose, identity and context of a given issue and its associated study. Without this, there is no study. Common elements are data that can be shared across a number of studies of similar nature — unit descriptions, speed of advance, system characteristics—items that don't change and yet can be used in a number of circumstances. These may be common within a specific mission area or level of warfare and therefore helps define and is composed of defined subordinate missions.

From a data development perspective, the core and common elements should be developed and available before the study. To the extent possible these can be prefabricated depending on the difficulty, importance, frequency of use and mutability of the data elements.



This graphic provides a picture of the concepts just discussed. Core elements define the scenarios and should be available for any study in the relevant area. The common elements are the data that can be shared across all studies in this arena. It includes environmental and geographic data and, for many studies, generalized threat laydowns and enemy objectives for a conflict. Core and common data should be developed centrally and provided to the study community. The custom data is that data that must be altered from a standard or produced uniquely to fit a specific study topic. The organizational structure and mission effectiveness of a new concept, for example, must be developed to determine the overall effectiveness of such a concept in the context of a larger war operation. Once a new concept is validated and adopted, it can become part of the next study's common data.

Policy Issues

		Present Condition		
		Institutional Provider	Executive Agent	Local Tribe
		for the DoD Enterprise	for the Comm Of Interest	for the Private Interest
	Core	Not Suitable		
Content	Common		Roll Your Own	
	Custom			"Un-Obtainium"

		Present Condition		
		Institutional Provider	Executive Agent	Local Tribe
		for the DoD Enterprise	for the Comm Of Interest	for the Private Interest
	Core	Not Suitable		
Content	Common		Roll Your Own	
	Custom			"Un-Obtainium"

The working group developed the graph above to suggest where data is currently developed while the next chart looks at how this might evolve through time. The present condition does not have a common provider for the core data and the type of data that should be common is normally produced by each organization as it conducts a study. As to custom data, this is generally not shared with any other organization and may be masked in studies provided to the DoD.

Policy Issues

		Desired Condition		
		Institutional Provider	Executive Agent	Local Tribe
		for the DoD Enterprise	for the Comm Of Interest	for the Private Interest
	Core	Immutable		
Content	Common	Baseline (D, I, F)	Assembled	
	Custom		Consensus	Tailored++



A desirable future would be for the DoD to provide the core and much of the common data while a consensus activity generates the remaining common data and much of the custom information. Only as required would individual organizations tailor the custom information and even then they would do that in a transparent way so others could employ their methodology and expect to obtain similar answers.

Way Forward

- **Immediate: Recommend these content baselines for the core**
 - Unit Order of Battle (UOB)
 - Mission Essential Task Lists (METL)
 - Weapon effects
 - Threat representations
- **Intermediate: “Joint” funding for core content using COI priorities**
 - Example: JMEM plus-up for CC requirements
 - Example: Proposed CTEIP funding of TMAP
 - Example: Possible JTIMS support for non-training METL's
- **Long Term**
 - Need a demonstration/experimentation of framework, procedures, value-added of top-down definition/stocking of core and common "data+algorithm for purpose in context" followed by bottom-up reconstitution of COI, private use.



The working group recommended core information provided by the DoD include the type of information show above. In addition, to provide an incentive to develop and use common data, the group recommended joint funding using community of interest priorities for specific classes of data. In the long term, the group recommended a demonstration/experiment to show the value of using the principles discussed here in the application of a study. Once the demonstration showed the value of having the community provide joint access core and common data to guide studies, this approach should be implemented throughout the DoD analytic community.

Improving Defense Analysis
Through Better Data Practices
March 25 - 27, 2003, Alexandria, VA

**MORS Workshop Outbrief Report
Working Group #3: Creating Data Internally**



**Chair: Dr. Dennis Leedom
Co-Chair: Bryan Paris**

Members

- Dr. Laurel Allender, ARL
- Mr. Donald Bates, JWARS
- Dr. Barry Bodt, ARL
- Ms. Emily Clark, DFI Int'l
- Mr. Sherman Cockrell, AFLMA
- Mr. Ezekiel Dennison, JDS-OSD
- Dr. Richard Fitts, OSDPA&E/JDS
- Mr. Robert Gangler, ARL
- Ms. Margaret Keish, NGIC
- Mr. John Kirzl, EBR, Inc.
- Mr. Jin Kwon, AMSAA
- Dr. Joe McDaniel, AFRL
- Lt Col Gregory McIntyre, OSDPA&E/JWARS
- Mr. Dale Fye, JFCOM J-9
- Mr. Scott Simpkins, JHU/APL
- Dr. Robert Smillie, NCCOSC
- Mr. Timothy Smith, ONI
- Mr. Jordan Wescott, ASC/ENMM



This slide shows the participants in Working Group 3.

WG Charter – Focus, Approach, Methodology

- **Problem Area**
 - Review best practices for generating data (operations, field tests/exercises, laboratory) and accrediting its use in M&S
 - Identify technical/organizational obstacles for data validation/accreditation
 - Identify data sources for each warfare/system level X data category
- **Scope**
 - Technical system performance through strategic/global levels
 - All data categories
- **Approach/Methodology**
 - Summary presentations
 - Group discussion



This is how we define the problem area, and how we approached developing insights and solutions.

Critical Properties of Data in M&S

- Data should be operationally defined in terms of relevant real-world phenomena and performance outcomes
- Data should be measured and recorded in a scientifically reliable, valid and consistent manner
- Data should be used in M&S in a manner that is consistent with its original definition and context
- Operational definitions are influenced by the maturity of our science and theories regarding warfare
- Data reliability and validity are influenced by the degree of experimental control and precision/standardization of our metrics
- Consistent use is influenced by the availability of formal mechanisms that provide accreditation and adaptation to specific analytic uses



This slide serves as a framework for our discussions. It outlines what we consider to be three critical properties of data and how it is used in M&S. The critical properties are shown on the left. Shown on the right are what we consider to be factors that influence our attaining these properties.

Technical Obstacles / Barriers

1. **Human/cognitive/organizational performance modeling remains an emerging area of science**
 - Qualitative assessment often substitute for quantitative data
 - Metrics constantly evolving – provides little basis for comparing findings over multiple studies/exercises
2. **Lack of a verified theory of war – needed to frame modeling and data requirements in transformation areas such as IO, effects-based operations, MOUT, peace enforcement, asymmetric warfare, etc**
 - Many transformation concepts lack operational definition
 - Face validity substituted for scientific validity
3. **Data aggregation often involves complex, non-linear processes that are not well understood**
 - Lack of metrics for complex, emergent processes
 - Lack of data standards for non-linear modeling (e.g., agent-based modeling, system effects)



In our WG we identified a number of obstacles and barriers to creating valid, reliable and relevant data. This slide summarizes those obstacles that were considered technical in nature.

Number one, we considered the human, cognitive, organizational performance modeling area to be challenging, and that science/data in this area has not yet matured to robust, quantifiable level. Specifically, the data generation in this area is inhibited by the lack of consistent, quantitative metrics.

Number two, defense transformation reflects number of warfare concepts that have not been fully developed in terms of theory and measurement.

Number three, we now realize that warfare involves a number of complex non-linear processes that are not well understood. This inhibits our ability to aggregate or link data from one level of warfare analysis to higher levels. These aggregation techniques and practices need to be captured and standardized.

Institutional Obstacles/Barriers

1. **Representation of human/cognitive/organizational performance lacks a Joint-Service organizing mechanism comparable to JTCG/ME model**
 - Impacts our ability to assess IO, effects-based operations, MOOT, peace enforcement, asymmetric warfare, etc
 - Many pockets of research, but no single organizing mechanism
 - Need to cover developmental systems and technologies
2. **Modeling VV&A generally remains an unfunded requirement**
3. **Limited dialog among researchers/analysts, model developers and operational experts**
 - Lack clearinghouse of up-to-date, comprehensive, authoritative data sources
 - Data requirements often lack detail and context
 - Lack of pedigree trail
 - Fear of misuse of data by other agencies/organizations
4. **Field tests, training exercises, and other operational venues often provide little basis for collecting data in a scientific manner**
 - Small sample size/single trial
 - Conflicting objectives (e.g., training vs evaluation vs experimentation)
 - Data capture availability/releaseability
 - Lack of standardized metrics / lack of trained observers



Likewise, the WG identified institutional obstacles shown here.

In the first area, we see a number of data areas (for example, human cognitive and organizational behavior) lack joint Service organizing mechanisms comparable to what exists today in the JTCG/ME for weapons effects.

In the second area, we generally note the lack of funding available for data and model VV&A.

In the third area, we note that there is limited dialogue existing among researchers/analysts, model developers and operational experts.

Finally, we note that creating data from field tests, training exercises and other operational venues is problematic for the reasons shown here.

[illegible]

91

The Road Ahead

1. **Develop family of canonical data standards within each category of data**
 - Forces/Units, CP&A, Scenarios/CONOPs, Log/Force Flow, C3I, Decision Behavior, Soft Factors/Training/Morale
 - Blue, red, green, brown
 - Needed for reuse of data by different customers
 - Requires authoritative voice across Services
2. **Recognize that data requirements and modeling paradigms are interdependent and evolve simultaneously**
3. **Articulate more fully the types of questions being addressed by M&S so that data providers have a better contextual understanding of how data will be used (*top-down approach*)**
4. **Provide feedback to data providers regarding areas of greatest sensitivity in M&S → leads to refinement of data creation priorities**
5. **Factor in the costs and benefits of requesting specific types of data for M&S**
6. **Develop methods (e.g., digital libraries) and incentives (e.g., charters, funding, data rights, controls on use) for data sharing across different communities (*willingness to share data and use data from other sources*)**
7. **Transformation requires fundamental relook at assessment methodology, functionality and analyst training/education**



The WG identified a number of steps that we feel could be taken to improve our collective ability to create valid, reliable and relevant data for M&S.

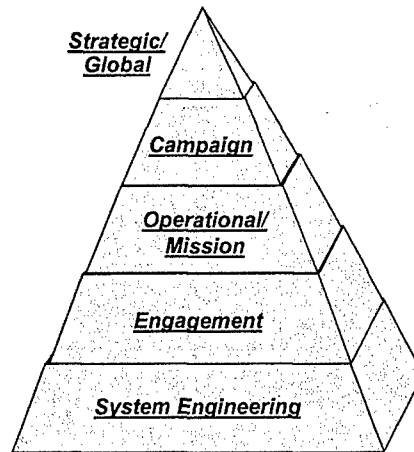
Promising “Best Practices”

1. **Apply AMSAA/JDS organizational concepts across other Services and system/warfare analysis levels**
 - Provides accreditation bridge between data providers and specific customers within M&S community (more than just a library or resource website)
 - Requires authority, funding and manpower authorization to be effective
 - JDS addresses only strategic/campaign analysis level → need to expand concept to lower warfare levels within M&S community
 - AMSAA addresses weapons effectiveness data → need to expand concept to logistics, C3, decision modeling, etc
 - OSD Analytic Agenda initiative only addresses strategic level → must rely on Service initiatives/funding to expand concept to other areas / levels
2. **Experimentation campaigns provide best mechanism for linking multiple levels of system/warfare performance**
3. **Analytic audit trail/traceability tags required for tracking pedigree of data**
 - Tags provide M&S customers with understanding of reliability, variability, relevance and contextual validity of data

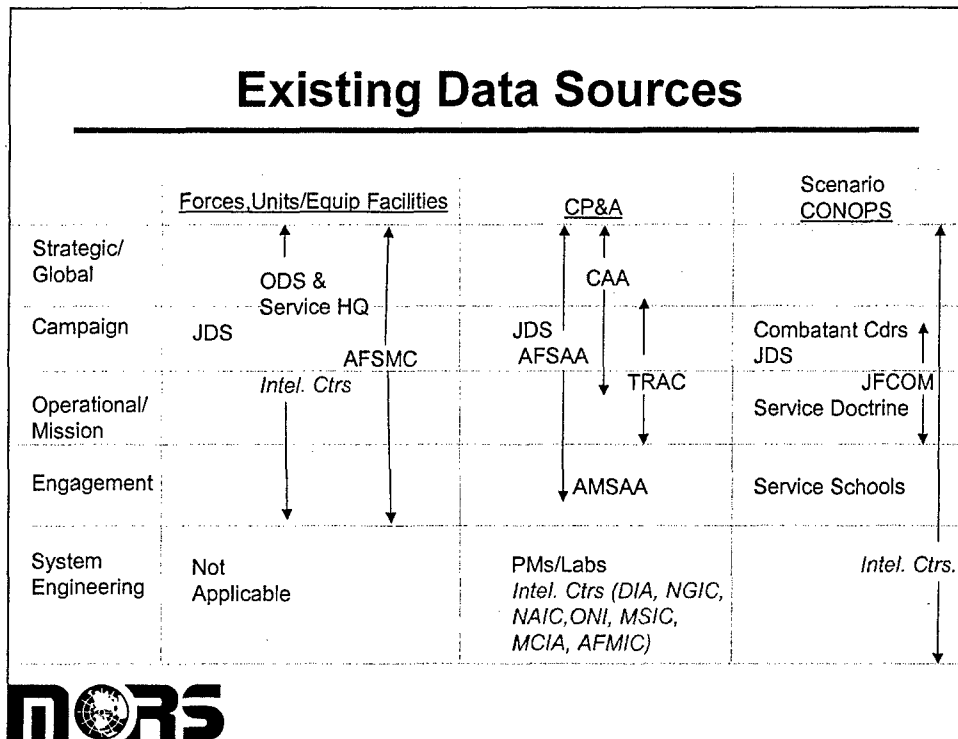


Next we identified promising “best practices” that could be generalized or extended across DoD.

Levels of Warfare Analysis

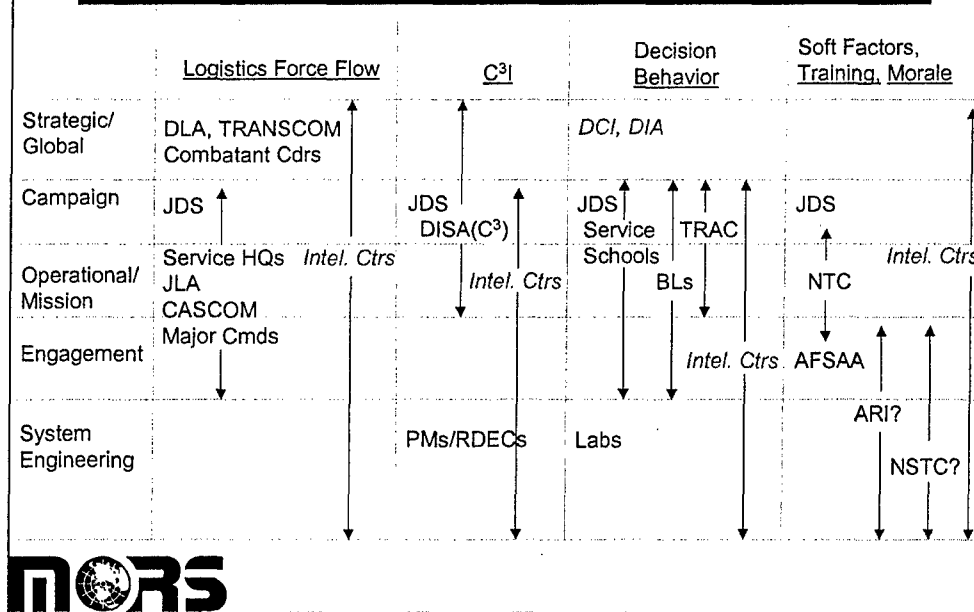


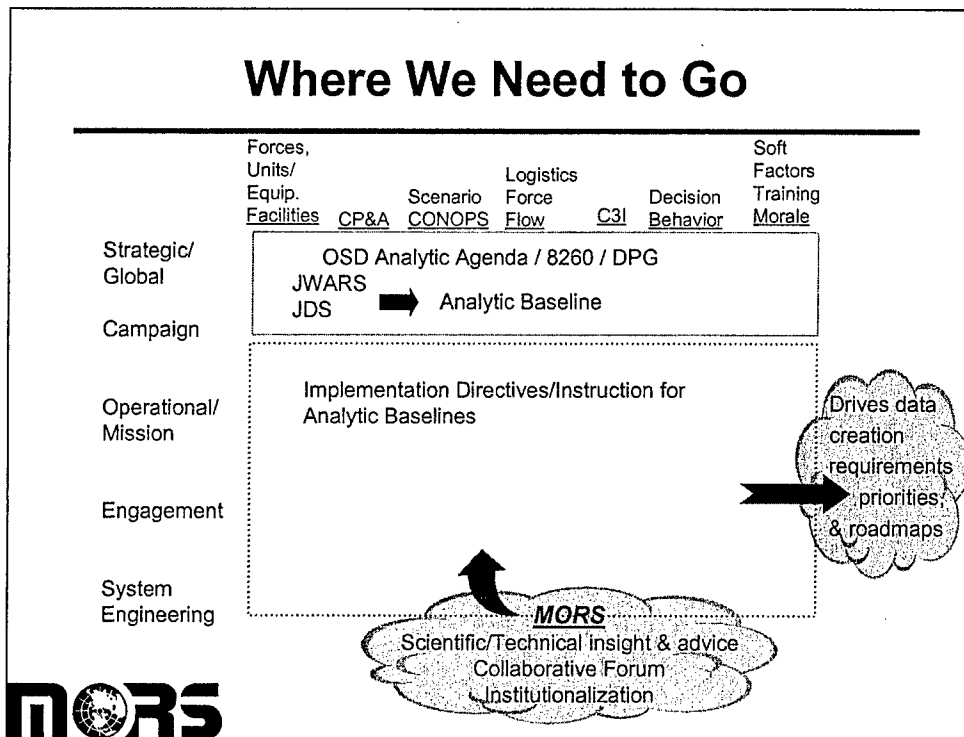
Before looking at existing data sources, we note here the levels of warfare analysis considered in our discussions.



This chart and the next examine each level of warfare analysis, and identify some of the organizations that provide data in the major categories relevant to M&S. Note, intelligence agencies that provide data on green, red, and brown systems are depicted in red (and are italicized).

Existing Data Sources





Our discussions concluded with an assessment of where we need to go in terms of organizational guidance and structure. The major point to be made with this chart is that current OSD analytic agenda extends only down to campaign level. We believe it is necessary for the Services to extend the concept and mechanisms of the analytic baselines down through each of the lower levels. We understand that there will be many Service checks and balances involved with this process. However, this process is necessary to provide the framework that will drive data creation requirements, priorities, and data creation roadmaps. Also depicted is the role that we believe MORS should continue to play in guiding and enabling this process.

Follow-on Issues for MORS

1. Details of data standards

- Nomenclature/operational definitions/categories
- Resolution/fidelity
- Formats/tagging/metadata

2. Linkage of modeling paradigms and data standards

3. Process for data requests



We believe this workshop made a valuable contribution in raising the visibility of a number of important issues. However, we feel more should be done in this regard. Shown here are a number of follow-up issues that could be addressed in future MORS meetings. We recommend that these issues be translated into appropriate Terms of Reference for these future meetings.

Improving Defense Analysis
Through Better Data Practices
March 25 - 27, 2003, Alexandria, VA

MORS Workshop Outbrief Report
Working Group #4: Converting and Employing Data



Chair: Clay Bowen
Co-Chair: Rudy Pabon
Co-Chair: Ernie Boehner

Working Group 4 addressed issues associated with the application of data.

Members

- Ernie Boehner
- Clay Bowen
- John Beilfuss
- Edward Brinko
- Paul Chang
- Richard Cobb
- Ty Coburn
- Donald Duckro
- Venton Duncan
- Louis Finch
- Robert Graeberner
- Don Hodge
- Philips Isaacs
- Susan Iwanski
- Kathleen Jackson
- Bruce Kind
- Geoff Koretsky
- Richard LaRiviere
- Rudy Pabon
- Michael Runnals
- Roy Scrudder
- Maj Thomas Sloan
- Mark Taylor
- William Troy
- Randall Wimmer
- Christopher Winkler



Clay Bowen, Rudy Pabon and Ernie Boehner led the group. Members represented a broad cross-section of the data collection and analytic study community.

WG 4 – Focus

Transforming and Employing Data. This “Data Application” group will view current DoD and best practices of data application, analysis and sharing to include transforming data for specific purposes and using data effectively. The WG will address appropriate use of data and maintaining data integrity when aggregating, disaggregating and transforming data for use in different applications. The WG will suggest ways to implement best practices as appropriate.



While the group had a broad charter, it focused on issues associated with data manipulation for use in models supporting operational and acquisition decisions. In particular, maintaining data integrity while aggregating, disaggregating or transforming it to fit the specific needs of DoD models or analytic approaches creates a series of challenges.

WG 4 – Approach

- **Define terms**
- **Focused discussions**
 - Data aggregation
 - Sharing data – security issues
 - Quality and appropriateness of data



The working group began by addressing terminology in order to ensure discussion was clear and meaningful. It followed with a number of presentations and focused discussions on aggregating data, security issues associated with sharing data with other groups, and finally on quality and the appropriate use of data for specific applications.

Discussion - Data Conversion

- **“Conversion” vs “transformation”**
- **Conversion trail**
 - Raw
 - Central management
 - Tool specific format
 - Analysis results
- **Two conversions**
 - Media
 - Language
- **The types of translation (format, mathematical, mapping and aggregation) are not independent. For example, some aggregations may be stated in terms of mathematical functions.**



In the terminology discussion, the group decided to change part of its name to better reflect its focus. Because “transformation” has taken on a specific set of meanings in the DoD, the group elected to call data manipulation for a specific analytic purpose “data conversion.” When converting data, it is essential to leave an audit trail so that others will know exactly where the raw data originated and how it was manipulated to fit the specific use. In the event of later questions or a future desire to extend the data to new applications, the audit trail is essential. The group agreed that data standards and use of metadata would be a major step forward; at the same time, the group felt it would be a mistake to consolidate military data under a single centrally managed repository. Not only would this be a major challenge to establish and run efficiently, it would not be responsive to the many requirements for data throughout the DoD and Services.

The group also noted that data drives analysis and that care must be taken to ensure the data used is appropriate for both the supporting tools as well as for the overall analytic purpose. Often raw data must undergo two conversions just for application and errors can be introduced in either conversion, between media types or in terms of language. Some models, for example, use a single value to represent a radar cross section of an object in space. Others require a full cross-sectional representation from all angles in 3 dimensions. In most cases there is no easy way to translate from one to the other. What values are used and what confidence can be attributed to model results using this data depends on the context of the study itself and the desired area of insight.

A main point of this is that types of translation are not independent and analysts and planners must use care in the conversion.

Definition

- **Data Conversion** – The translation of data from one representation format to another including modifications of semantics (meaning) and/or syntax (storage representation). Data conversions include:
 - Format – changing the “bits and bytes” arrangement of data for specific application use.
 - Mathematical – translation of data using a function which can be expressed as an algorithm. This embodies the mathematical definition of “function” where a set of inputs give one unambiguous output.
 - Mapping – translation from one discrete set of data values to another set of values by either objective rules or subjective judgments.
 - Aggregation – translation from a granular set of data values to a more abstract set of values.



The group developed the above definitions to be better informed during the ensuing discussions.

Definition

- **Data Sharing:** The process of providing existing analytic data among the levels of an organization, different agencies or services
- **Issues/problems/barriers**
 - Proprietary nature data
 - Bureaucracy
 - Lack of confidence
 - Timeliness
 - Regulations/laws
 - Fear of inappropriate use
 - Security or lack of security



Data sharing was seen as a critical aspect in successfully conducting cross function and cross-service studies. A number of barriers currently impede data sharing and make analysis intended to inform senior decision makers more difficult to conduct. The DoD would be well served to address these issues systematically in order to facilitate better application of data in all areas.

Definition

- **Security of data: Providing the appropriate level of protection to data and information to protect from unauthorized access and unauthorized changes**
 - Accomplished by
 - DoD classification levels
 - Access Control
 - Other protective labeling and controls
 - Privacy act data
 - Proprietary
 - Includes the required systems, procedures and training



Data security is important, but all too often security has been used as an excuse to preclude data-sharing between organizations. Data must be protected and must not be converted inappropriately. By the same token, data must be made available in the appropriate formats to allow study across a range of applications.

Definition

Scenario: An account or synopsis of a projected course of action or events. For purposes of this Directive, the focus of scenarios is on strategic and operational levels of warfare. Scenarios include information such as politic-military contexts and/or backgrounds, assumptions, operational objectives (threat and friendly), major force arrivals, and planning considerations. (DoD Directive 8260.1, Dec 6, 02)

scenario: Set of data and information needed to assess the capability of a force or force element to accomplish an objective.

ISSUE: Big "S" scenario definition is clearly for the strategic level. Small "s" definition allows for inclusion of mission focused analysis.



Scenarios are vital for the understanding of system and organizational performance or the measuring of capability in context. Scenarios themselves incorporate important data about the types of capabilities of interest to decision makers, the environments where systems must operate and the expectations about time, distance and adversary behavior that must be addressed by US military capabilities.

Data Storage and Retrieval

- **Current Department approach**
 - A wide variety of data storage schemes exist throughout the community
 - One size fits all approach is not appropriate
 - Lack of standardized formats and inadequate metadata leads to inappropriate use of data
- **AFSAA has developed a S&R system that could solve numerous agency data problems**
 - Injector tools can be written based on documented format of producer data and standard format for data repository
 - Application specific “ejector tools” can be written by the data consumer
 - Standard algorithms must accompany standard data

“To make it happen, management has to force structure on the analyst.”



The group felt a monolithic approach to data storage and retrieval would be a mistake and could make data development, conversion and sharing even more difficult. However, aspects of the data storage problem could be improved through the use of standardized formats and robust metadata. The Air Force Studies and Analyses Agency (AFSAA), among others, has developed an approach that has helped them address data problems internally and provide a better face to the customers who use AFSAA data for their work.

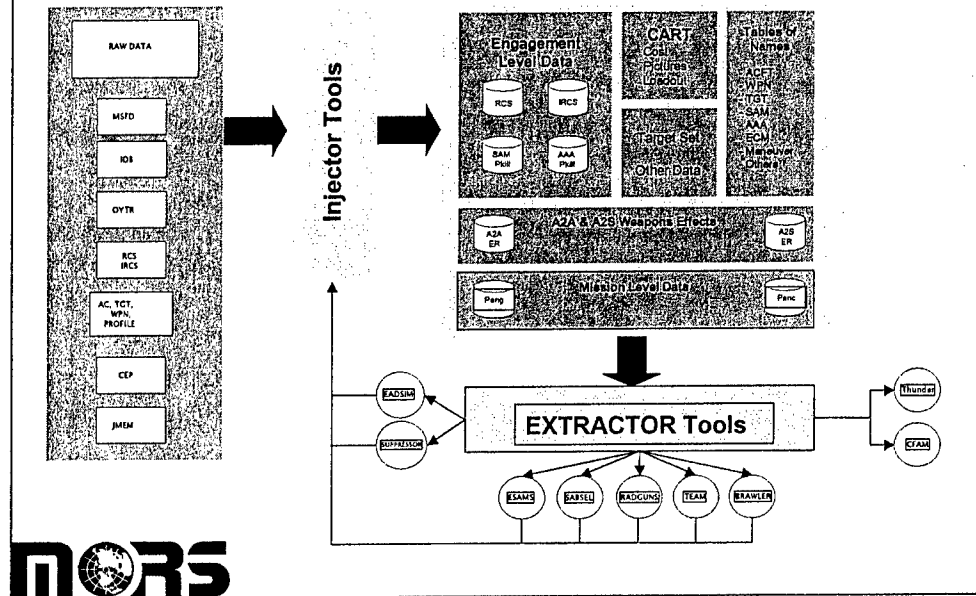
Data Storage and Retrieval

- **No one agency has its internal data house completely in order.**
- **Internal organization is the first essential step.**
- **Once analysis agencies have internal structure and standards (e.g. metadata) they can begin to share data.**
- **While the analysis agencies are developing their internal standards the DoD community should establish authoritative sources and rules for sharing data.**

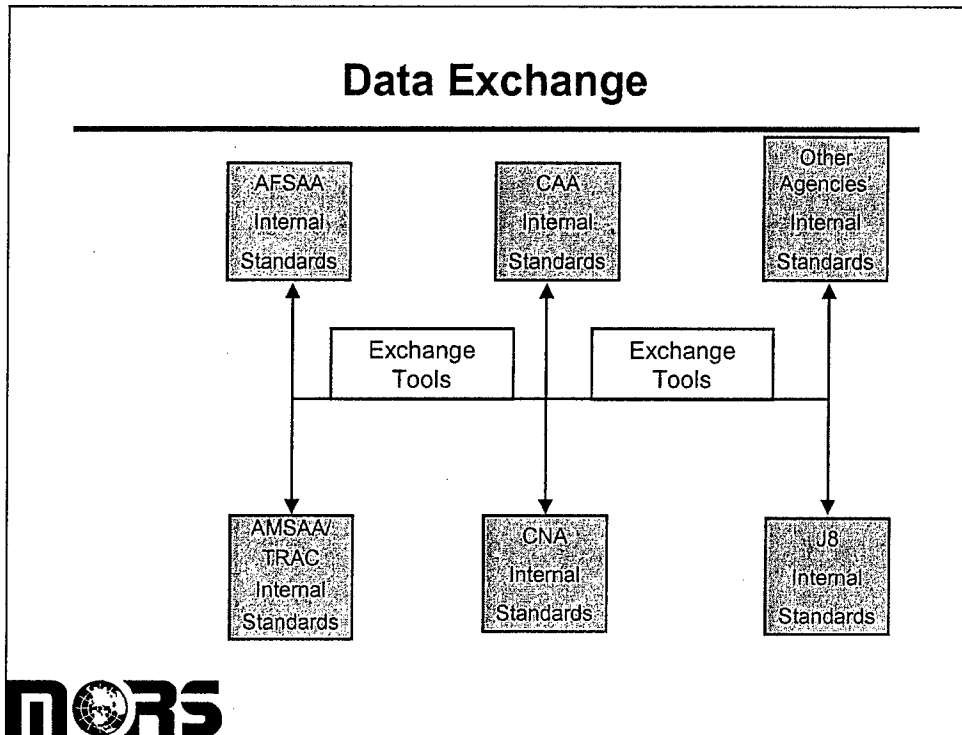


Although the group identified a number of organizations that could point to successes associated with internal data management, they concluded that no agency has its internal data house completely in order. The group felt that the first step would be for each organization to develop and apply internal structure and standards that could then be shared with others. This would provide the stepping stone for better data sharing. In the meantime, DoD needs to designate authoritative sources for data in various areas and evolve appropriate rules for the sharing of data.

Data Storage and Retrieval



The above graphic provides one view as to how a successful data storage and retrieval system might work.



Once major analysis organizations have in place the appropriate storage and retrieval capabilities, to include standards and metadata to describe data elements, these organizations can develop tools to exchange data while maintaining the integrity of the data elements. These can then be promulgated across the community to better facilitate data sharing.

Data Sharing

- **Technical issue:**

- How do we share in a way that everyone gets appropriate data for their studies?
- Technical issues are solvable

- **Trust issue:**

- Fear that data will be misused
- Fear that I'm being asked to "contract" for things I'm simply making a guess at



The bottom line is that the technical issues associated with data sharing are solvable. The cultural or trust issue needs work.

Trust Issues

- **Fear of data misuse**
 - Metadata inadequacies
 - Data storage and retrieval schemes
- **Fear of “contracting”**
 - Analytic baselines
 - JDS as honest broker



Within the cultural area, trust is the major barrier. Two broad categories can be used to classify why individuals worry about data they've developed being used or misused by others. One area has to do with the fear of data misuse.

Returning to the radar cross-section issue, passing a full 3-D radar return picture for all relevant radar frequencies could inadvertently be misinterpreted by another analyst in another organization who does not understand how the data was developed or how it is most appropriately applied. By the same token, changes in data storage and retrieval could inadvertently change data elements leading to erroneous results.

Likewise, understanding someone else's analytic focus and helping to convert data for their use could become a full time job, detracting from an organization's ability to conduct their own studies. The Joint Data System (JDS) can help in this by maintaining valid data for joint models that has been converted appropriately and is updated regularly to ensure some of these issues are avoided. Peer review is another way to ensure the appropriate use of data, but again such an approach imposes a resource requirement.

Quality and Appropriateness

- **Appropriate use of data is the responsibility of the data user**
- **The data provider has the responsibility to document assumptions that went into data development**



In the end, the group determined that appropriate use of data is the responsibility of the data user, while the data provider has the responsibility to document assumptions that went into the data development. Both the user and the provider need to keep the lines of communication open and to foster a more trusting environment.

Follow-on Activities

- Recommend a JDS workshop
- Recommend a MORS sponsored survey



To help in this area, the group recommended that JDS sponsor a workshop to both discuss their capabilities and clarify their needs. Most analytic organizations could learn from their internal procedures and a workshop could also ensure that while data is developed for internal studies, it can be shaped and prepared for joint community needs as well. In addition, MORS can continue to help lead in this area by surveying its membership for useful solutions to data conversion issues.

Survey

Characterizing the Investment in Analysis

Background: There is a common understanding that the analysis process is time consuming, data gathering and preparation consumes a significant portion of the time and resources, and there is a desire to shorten this process if possible.

Objective: We are seeking to understand where the time and resources are spent in the analysis process with respect to data and what barriers are possibly causing delay.

Survey Audience: We are seeking responses from those that conduct and manage analyses.

For your most recent analysis please characterize the following:

Phase	Duration (months)	Effort (person- months)
Scoping – Determining analysis objectives and establishing data requirements		
Data gathering – Obtaining data from external sources and from existing internal archives.		
Data conversion – Translating the data from the form in which it was received into the form needed for analysis. This includes running "lower level" models to derive data for the end-analysis tools.		
Analysis – Conducting the analysis, analyzing results, and preparing reports.		



The slide above provides an example of the type of survey that could be used to document the current time and effort required to generate data for major Service and DoD study efforts. Such a survey could help highlight for the leadership the current investment in these activities as motivation for adopting better standards and more open data sharing techniques.

Survey

- **What barriers (if any) caused time delays during each phase?**
 - Scoping
 - Data gathering
 - Data conversion
 - Analysis
- **Were you able to share the data used for analysis and the analysis results? If not, what prohibits this being done (e.g., organizational policy, time, resources, security, ...)?**
- **Where are data used for analysis and analysis results archived?**
- **Identification (optional)**
 - Analysis for which this survey was prepared:
 - Analysis organization:
 - Name:
 - Email:
 - Phone:
 - May contact you: yes/no



Here are additional questions that could be included in the survey.

International Note

Within UK MOD -

- **Director General (Scrutiny and Analysis)**
- **NOT part of acquisition chain**
- **Civilian component of MOD**

Acquisition programs must pass the DG(S&A) test



As a final comment, Working Group 4 benefited from international participation by a member of the United Kingdom's Office of the Director General for Scrutiny and Analysis. This office is not part of the internal Ministry of Defence acquisition chain, but it is required to comment on the quality and usefulness of analyses forwarded to government decision makers regarding alternative acquisition choices. Because fewer organizations are involved, data standards are easier to maintain, but UK analysts face many of the same issues confronted by their counterparts here in the US. In the end, however, having a single objective group review data and methodology has helped establish basic data standards to which all branches adhere to in the development of their analyses.

Improving Defense Analysis
Through Better Data Practices
March 25 - 27, 2003, Alexandria, VA

**MORS Workshop Outbrief Report
Working Group #5: Technology Support**



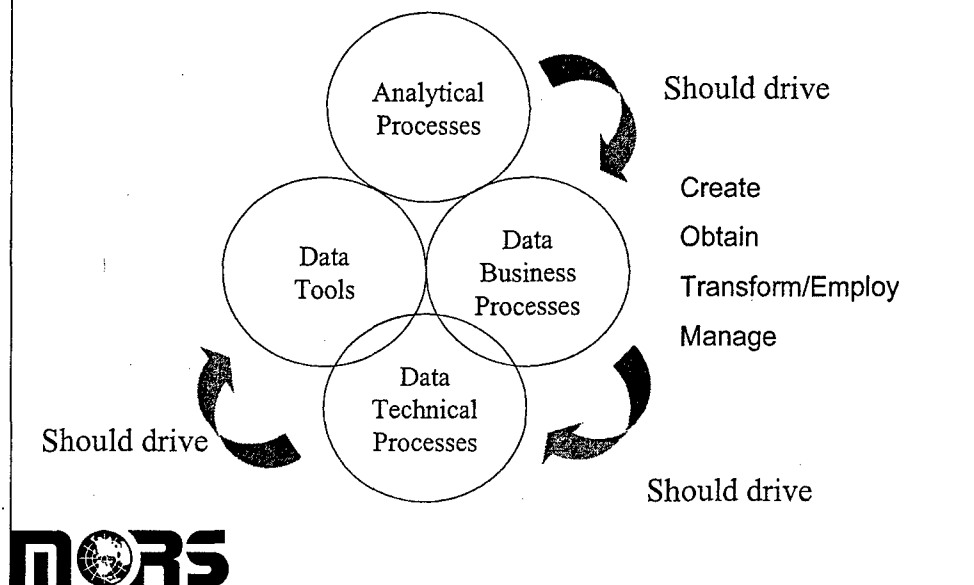
**Co-Chair: Bob Might
Co-Chair: Ron Smits**

Members

- John V. Harmon
- Rosemary Enright
- Robbie Williams
- Franco Tao
- Norm Burroughs
- Stephen Hale
- Dee Blodgett
- Bob Orlov (Spy and Synthesis)
- Bob Might
- Ron Smits
- Geoff Sherwood
- Lee Lacy
- Leslie Kann

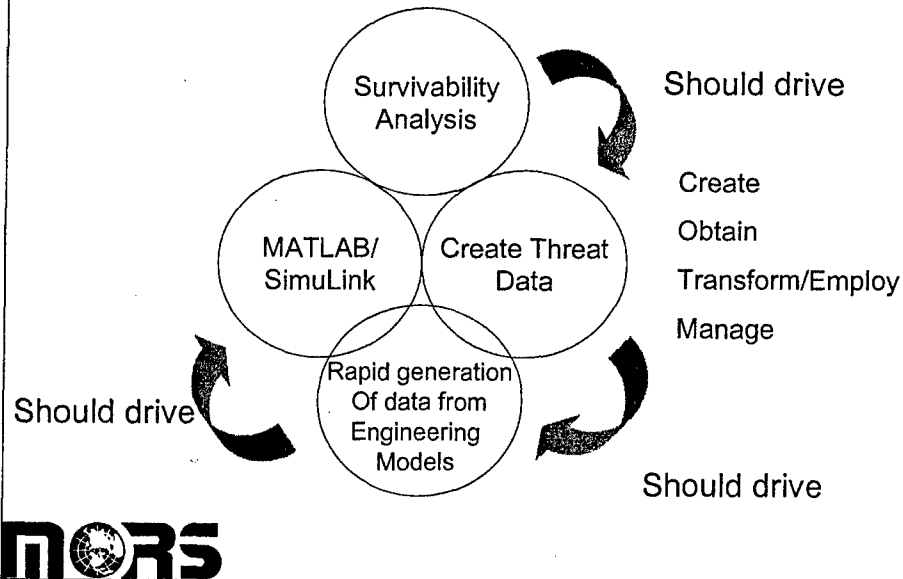


Assumptions 1



Analytical processes should drive data business processes, i.e., manage, obtain, create, transform. Data business processes drive technical processes. The technical processes should drive the data tools.

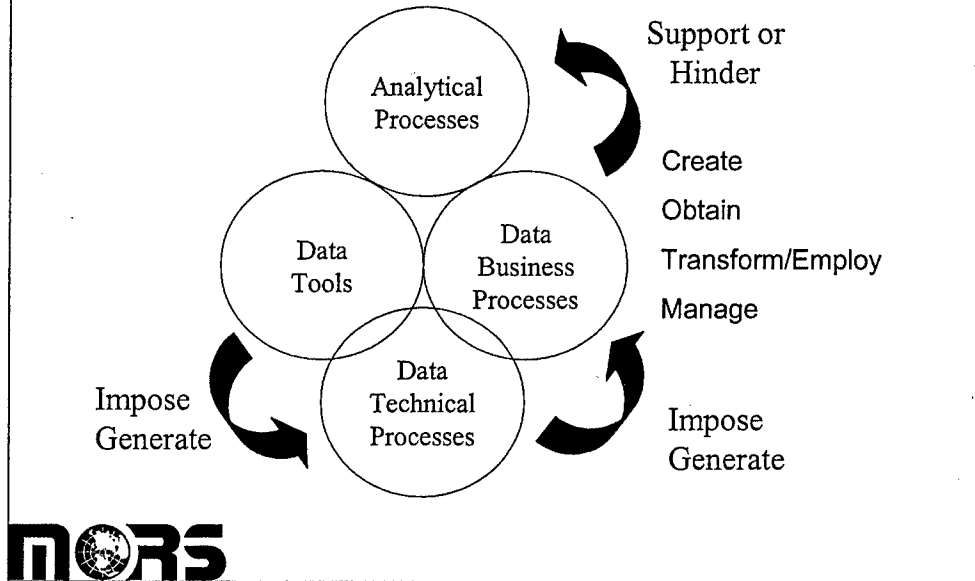
Example



This is an example to clarify these terms.

There is not always a clear distinction between the Business Process and the Technical Process. This is partially because the Technical Process can drive the Business Process.

Assumptions 2



Unfortunately, in many cases, the existence of a data tool will impose a technical process. Or when we are developing new tools we are formulating new technical processes. These technical processes in turn impose data business processes that may or may not be useful in the analytical process that they support. In many cases, they may actual hinder the analytical process.

Assumptions

- **The working group focused on the analytical data business process, not operational data creation**
 - Recognize that analytical requirements do not drive operational data creation
- **Postulated business process requirements that would come from other working groups**
- **Some technology solutions do not apply to all technical processes**
- **Business models are in different phases of life cycle**
 - Not in the same state of readiness



This slide is self explanatory.

Methodology Applied

- 1. Listened to briefs on current/emerging technologies**
- 2. Identified issues with existing tools and technologies**
- 3. Identified limited set of technical processes of interest**
- 4. Mapped technical processes to business processes**
- 5. Identified technologies needing support/development**



These are the five steps that we followed during the course of the workshop.

Briefings

- Semantic Web – Lee Lacy, Dynamics Research Corp
- Thread Model Assessment Process – Ollen Landrum , NAIC
- Data Integration Practices – Joe George, Data Junction
- Common Warehouse Metamodel – Geoff Sherwood, Adaptive Inc.
- XML tagging of unstructured data – Mike Bronder, CambridgeDocs
- JSF Authoritative System Database – Steven Hix, Paradigm Technology
- WebTAS Data Visualization – Brad Carman, ISS
- High-Dimension Data – Leslie Kann, CSI



Business Processes to be Supported

- **Create/Share Data**
- **Obtain Relevant Data**
- **Transform/Employ Data**



Concurrent development with the other working groups caused us to stay with high level requirements. We chose these three to pursue.

Create/Share the Data

- **Standardize the generation of metadata to include**
 - Stewardship/ownership
 - Security
 - Annotation on field by field basis
 - Metrics on both usage and quality
 - Versioning
 - Context of data generation and transformations
- **Automate the generation of metadata**
 - Support Meta Object Facility (MOF) standard
 - Automate ontology generation



One of the recurring themes from the briefings was the need to standardize the generation of metadata.

Create/Share the Data

- **Rapidly generate/collect and post meaningful data**
 - Automate data integrity checks
 - Error checking, consistency, configuration management
 - Propagate development of self documenting engineering models for rapid data creation
 - For example, Threat Model Assessment Process (TMAP)
- **Register Metadata**
 - Investigate standards such as Common Warehouse Metamodel (CWM)
- **Improve access control**
 - Standardize on one control system, such as Public Key Infrastructure (PKI), Common Access Card (CAC), Security Assertion Markup Language (SAML), etc.



Obtain Relevant Data

- **Improve data registry**
 - Investigate standards such as CWM
- **Enhance data discovery**
 - Make web content machine understandable (inferencing)
 - Support development such as Semantic Web
 - Support Ontology Web Language (OWL)
 - Support Metadata search engine development



Obtain Relevant Data

- **Integrate data from different sources**
 - Automate (to the greatest extent) extraction of data from sources with different formats
 - Encourage COI common data models
 - Continue to apply data interchange standards, such as XML
 - Support R&D of tools that adapt to changing data sources
- **Data mining (find interesting and actionable data)**
 - See output from MORS Data Mining Conference

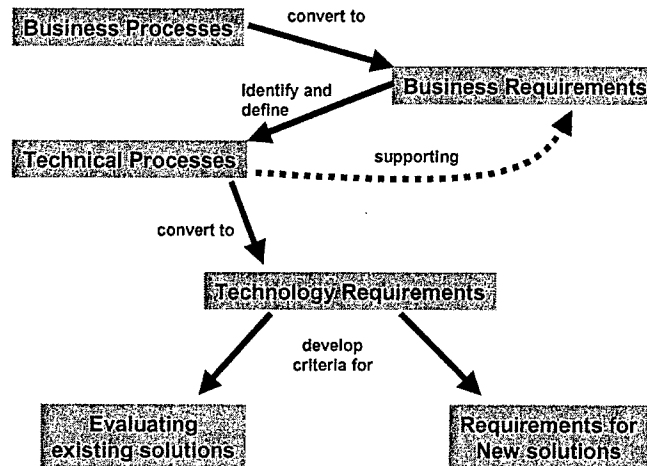


Transform/Employ Data

- **Improve data reduction methods**
 - Leverage existing methods, such as On-line Analytical Processing (OLAP)
 - Improve user-friendly data visualization/graphical tools
- **Extract meaningful data from unstructured sources**
 - Explore methods for characterizing unstructured sources to extract meaningful data
 - Evaluate cost/benefit of unstructured extraction



Recommended Technology Development Process



Summary

- A combination of data tools and business requirements drive our existing technical processes
- COTS and GOTS tools exist that enable the current technical processes. These have been developed ad hoc and backfit to the business processes
- In the future, data business processes will be different. These new processes need to control and drive the development of tools



Technology Development Process

- Identify business processes to be supported
- Convert business process to requirements
- Identify and define technical processes that support business process requirements
- Convert technical processes to technology requirements
- Based on technology requirements, develop the criteria to
 - Evaluate existing technical and non-technical solutions
 - Develop requirements for new solutions



Sample criteria: Easy, fast, cheap, good, interfaces, protocols and formats, expose metadata effectively, scaleable, central integration, self documenting.

Because we were working on the final steps without the first steps being completed, what we have done is a partial example of what should be done. This is a process that we believe needs to be done in the near term – 6 to 12 months – to ensure positive return on investment. We believe that the other 4 groups should have made a start on step 1.

Discovery/Observations

“One Size DOES NOT fit all”

A technical process that supports a data business process for one type or category of data does not necessarily work for other data types or categories

Conclusion: We need some schema for categorizing data to link technical processes to business processes and to identify their relevance to the business process



Technologies Discussed/Briefed

- **Semantic web technologies**
 - Enhance discovery process
 - Link ontologies
 - Enable semantic joins
 - Allow machines to read and inference about web content
 - Represent data in XML/RDF/OWL
 - ISSUE: Scalability for inferencing
 - ISSUE: Ontology registry required



Technologies Discussed/Briefed

- **Modular engineering models**
 - Streamline data creation
 - To rapidly populate databases
 - Empower data end user
 - Respond to changing user needs
 - Capture/output of anything calculated
 - Use model to generate specific data set automatically
 - Embed in other simulations
 - Reusable
 - Modular
 - Hyperlink to engineering descriptions
 - Standardize on MatLab/Simulink
 - ISSUE: Configuration management/data integrity



Technologies Discussed/Briefed

- **Distributed application integration**
 - Perform data normalization
 - Perform data validation
 - Provide capability to enrich data
 - Emerging technologies
 - Improve visual design tools
 - Increase metadata transparency
 - Standardize XML data definitions
 - Implement more efficient web services
 - Automate access to web sites to extract and insert data
 - Perform graphic recognition
 - ISSUE: History of data use/reuse
 - ISSUE: Track of data generation/transformation



Technologies Discussed/Briefed

- **Common Warehouse Metamodel**
 - Continue improvement and implementation of standards
 - UML
 - XMI
 - MOF
 - CWM
 - Support
 - Stewardship/ownership
 - Browse and update
 - Security
 - Flexible categorization/ annotation on field by field basis
 - Search/reuse
 - Collaborative working
 - Notification
 - Metrics both on usage and quality
 - Versioning
 - Federation of catalogs
- Track transformation of data
- ISSUE: Tracking data provenance
- ISSUE: Provide a more efficient search engine for Metadata
- ISSUE: Advanced automated feature selection
- ISSUE: Ground truth, temporal tagging
- ISSUE: Data reduction



Technologies Discussed/Briefed

- **Unstructured content to meaningful XML**
 - Implement portals
 - Use tools for creating XML for content as produced
 - Apply methods for automating XML for legacy documents
 - Learning based extraction
 - Rules based extraction
 - Pattern recognition
 - ISSUE: Natural language processing
 - Semantic processing
 - Computational linguistics



Technologies Discussed/Briefed

- **Authoritative Systems Database**
 - Create flexible extensible architecture for rapid update
 - Analyze legacy data
 - Support transfer to common semantics and syntax
 - Create custom designed user interfaces for unique applications
 - Correlate attributes across simulation levels/systems
 - Join data from multiple sources
 - Used IDEF1X for data model
 - Automate XML/XSD/XSLT creation
 - Data requirements in coherent rule-based structure
 - Mapping of metadata
 - Enforce data consistency
 - Ensure use of the most current data
 - Better ways to integrate and organize data requirements for multiple simulations



Technologies Discussed/Briefed

- **WebTAS**
 - Visualized relational, flat, XML
 - Hyperlink records
 - Model/time event detection
 - Solutions sent out via web pages
 - XML import/export



Technologies Discussed/Briefed

- **High-Dimension Problems**
 - Machine learning
 - Discovering patterns of information
 - Feature extraction/enhancement
 - State machines
 - Natural language processing
 - Criteria
 - Interestingness
 - Novelty
 - Requesting tools for related information
 - Tools for analyst to decide what he needs to look for
 - Cognitive agents
 - Need ground truth, temporal tagging
 - Support: prioritization, data/event relationship, selection of data/event retrieval, predictive, how can I impact



Improving Defense Analysis
Through Better Data Practices
March 25 - 27, 2003, Alexandria, VA

**MORS Workshop Outbrief
Synthesis Panel Report**



This section contains the report of the Synthesis Panel for the MORS Workshop “Improving Defense Analysis Through Better Data Practices,” conducted at the Institute for Defense Analyses, 23-25 March 2003.

Lest We Forget ... Context Matters



©Washington Post, March 27, 2003



Agenda

- **Panel Goals, Composition**
- **Insights on Nature of the Problem**
- **Selected Findings, Recommendations**
- **Summary**



The Synthesis Panel report consists of four sections.

As a context, the first section identifies the goals, objectives and composition of the Synthesis Panel.

The second section summarizes insights on the nature of the problem that the Synthesis Panel derived. These insights were developed from the remarks of the plenary speakers, the internal discussions of the Synthesis Panel members, and the deliberations of the other six panels.

The third section formulates selected findings and recommendations based on the workshop deliberations.

The final section briefly summarizes key observations and conclusions.

Panel Goals, Objectives, Scope

- **Goals**

- Provide an overview of the entire workshop

- **Objectives**

- Clarify the nature of the problem by conducting internal panel discussions
 - Capture the state-of-the-practice based on the presentations at the Plenary
 - Derive key findings and recommendations (including best practices) from the results of the individual panels

- **Scope**

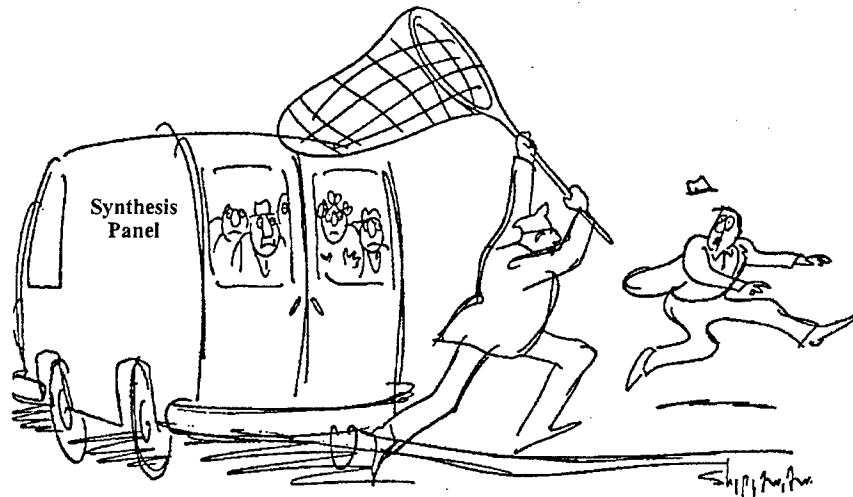
- Focus on data to support *strategic analyses*
 - Consideration given to data to support analyses for acquisition, OT&E, operations



The Synthesis Panel had one major goal: to develop a better understanding of how to improve defense analysis through better data practices, from a holistic perspective.

Consistent with that goal, the Synthesis Panel pursued three supporting objectives. First, it sought to clarify the nature of the problem, based on internal panel discussions. Second, it sought to capture the state-of-the-practice in data practices, based on the presentations at the plenary. Finally, it derived key findings and recommendations based on an integration across the results of the individual panels.

Recruiting Was Tough...



"But I don't want to work the data problem!"



As many of you know, members of the Synthesis Panel have a “daytime job” and a “night-time job.” During the day, each member of the panel is assigned to one of the five mission oriented panels. During off-hours we meet to share insights and develop a holistic view of the subject. Thus, recruiting can be difficult, requiring unorthodox approaches.

Panel Composition

- Phil Barry (MITRE) [WG-2]
- Mike Hopkins (Computing Technologies) [WG 1]
- Bob Orlov (JS, J-8) [WG-5]
- Roy Reiss (USAF) [WG-3]
- Cy Staniec (Northrop Grumman) [WG-1]
- Stuart Starr, FS (MITRE) [Floater]
- James Stevens (PA&E) [Floater]
- Gene Visco, FS (Visco Consulting) [WG-2]
- Kirk Yost (L-3 Com Analytics) [WG-4]



The above list identifies those individuals who ultimately “volunteered” to participate on the Synthesis Panel. As noted above, each of them also participated in the deliberations of one of the other five workshop panels.

The following representation was provided on the panel:

- Government-3
- FFRDCs-2
- Private industry-4

Agenda

- Panel Goals, Composition
- Insights on Nature of the Problem
- Selected Findings, Recommendations
- Summary



The Synthesis Panel developed insights on the nature of the analysis — data problem by taking advantage of several sources. This section briefly summarizes those findings in the following areas:

- Broad insights developed by the Synthesis Panel through its panel deliberations.
- Insights acquired through the presentations at the plenary. These insights were aggregated into three categories: 1) insights from DoD; 2) insights from non-DoD agencies; and, 3) cross-cutting insights.
- Insights acquired and captured in the other panel discussions. The Synthesis Panel assembled several insights that were identified in one or more of these panels.

Background: The Problem is *Not* New (1 of 2)

"The government are very keen on amassing statistics.

**They collect them, add them, raise them to the n-th power,
take the cube root and prepare wonderful diagrams.**

**But you must never forget that every one of these figures
comes in the first instance from the village watchman, *who
just puts down what he damn pleases.*"**

**--Comment of an English judge on the subject of Indian
statistics; Quoted in Sir Josiah Stamp in "Some Economic
Matters in Modern Life"**



Background: The Problem is *Not* New (2 of 2)

- “Data! Data! Data!” he cried impatiently. “I can’t make bricks without clay.”

Sherlock Holmes

- “Without data we are nothing!”

Walt LaBerge, SIMTECH 1997

- “Theory without data = philosophy; data without theory = noise”

Anonymous



Perspectives Presented at the Plenary

- **DoD perspectives**

- Data enterprise
- Analytic agenda
- Data Management (DM)
- Data directive and implementation
- Panel: Services, Joint Staff

- **Perspectives of Non-DoD agencies**

- State
- NOAA
- NSF
- NIH



During the Mini-Symposium that initiated the event a distinguished group of four speakers, followed by a DoD panel and additional speakers from non-DoD agencies, addressed the subject of data goals in DoD and the other agencies. The following section briefly summarizes selected perspectives on those issues.

Plenary: Views of Data Goals

- Throughout the plenary presentations, views of *data goals* were presented that are similar but subtly different; e.g.,
 - John Osterholz on Enterprise Data Strategy Goals
 - AVAILABLE: Visible, accessible, institutionalized
 - Use: Understandable, trusted, interoperable, responsive to user needs
 - Charles Swett on the DoD Analytic Agenda
 - Transparent, RAPIDLY available, accurate, standardized, sufficiently detailed, synchronized with PPBS
 - Tony Simon on Transforming the Way DoD Manages Data
 - Ensure that all data are visible, available, useable when, where needed
 - Make data visible, accessible and understandable across DoD and beyond
 - Jim Stevens on DoD Data Directive...
 - Visible, accessible and traceable data ready to support analysis



Plenary Major Points: John Osterholz

- **Standardized data element effort of the 1990's**
 - “A collapsed, failed effort”
 - “It taught us that there is an important sociological component to the problem”
- **Paradigm shift: TPED → TPPU**
- **“Data fuels many value chains; it may be intended for one, but be valuable for many”**
- **Major elements, trends of Enterprise Data Strategy**
 - Private data (DOWN)
 - Community of Interest Data (UP)
 - Enterprise Data (UP)
 - Data Mobility (UP)
- **“Interoperability begins at the data level, not the systems level, has a common language, common tagging. Data is our most important product.” (Brig Gen Mike Ennis, USMC Director of Intelligence)**



Plenary Major Points: Charles Swett

- **“As DoD transforms, the *analytic system* must transform”**
 - Perform analyses that are higher quality, better synchronized, more flexible, responsive, able to cope with unanticipated needs
 - Deal with a wider range of variation/future uncertainties
- **Transforming the *joint data* enterprise is just one element of the analytical transformation; it also requires**
 - Capabilities-based planning
 - Joint operational concepts
 - Wider array of approved tools (e.g., deal with IO, C4ISR)
 - Enhanced analytic resources
- **Key products: *analytical baselines* to provide visible, accessible and traceable data ready to support Department-level analyses**



Plenary Major Points: Tony Simon

- **Key barriers for the *users* of data**
 - Cannot (or does not) communicate with the data producer
 - Unaware that data exists
 - Cannot access data (due to organizational or technical barriers)
 - Cannot use it (due to lack of understanding of what the data represents)
- **Thoughts on overcoming barriers (*shifting power to the user*)**
 - Establish Communities of Interest
 - Implement DoD Discovery Metadata Standard (DDMS)
 - Guide Data Management activities through user-driven metrics, processes
 - Provide infrastructure, services to permit user to find, retrieve data



Plenary Major Points: Jim Stevens

- **Summarized DoDD 8260.1 and DoDI 8260.2**
 - Establishes responsibilities and procedures (i.e., Policy, JS, PA&E, DoD components)
 - Creates Joint Analytic Data Management Steering Committee (JADMSC)
 - Characterizes an “Analytical Baseline” (warm databases ready to support Department-level analyses); a package comprising
 - A scenario
 - Concept of operations
 - Integrated data
- **Goal: Visible, accessible and traceable data ready to support analysis**



Plenary: Service, Joint Staff Perspectives (1 of 3)

- **Jackie Henningsen, FS, AFSAA**

- Identified key USAF drivers; e.g., analysis needs growing (timelines for QDR, budget, BRAC); Government analyst numbers shrinking
- Outlined evolving CONOPs analysis process

- **Vern Bettencourt, FS, HQDA/G3**

- Discussed extending AMSAA focus to developmental systems, Army data initiatives (e.g., FCS Advanced Collaborative Environment; Army data support to JWARS)
- Identified residual issues; e.g., level of preprocessing of data required for aggregate models; reliance on unclassified data; conflicting data among models within federations; capturing training and operational data)



Plenary: Service, Joint Staff Perspectives (2 of 3)

- **Greg Melcher, N-81**

- Described collaborative effort to develop USN/USMC baseline for "Win Decisive" analysis
- Exploring "difficult to quantify" areas (e.g., FORCEnet, anti-access strategies)

- **George Akst, MCCDC**

- Provided a case study on OEF
- Challenges included Data Management (DM), security glitches, limited USMC lessons learned system, making system interactive
- Lessons learned areas to plan for include file structure, DM, choice of database, security



Plenary: Service, Joint Staff Perspectives (3 of 3)

- **Pete Byrne, Kevin Kelley; Joint Staff**

- Joint Staff observations and cautions on the
 - Characteristic timescales of data
 - Need to associate algorithms with data
- Provided a case study based on the Operational Availability study; challenges included
 - Inconsistency of data from Services (substance, format)
 - Need for better processes (e.g., verify data in a time dynamic environment)



Non-DoD Insights (1 of 2)

- **Humanitarian Information Unit (HIU) -- DoD's data problem of today and the future?**
 - Multi-sectoral (e.g., political military, health, shelter)
 - Multi-dimensional (e.g., affected population, vulnerable groups, assistance)
 - Contextual (e.g., historical, cultural, ethical)
 - Multi-source (e.g., government, UN, NGOs, media)
 - Non-standardized (e.g., formats, definitions, indicators, measurement indicators, methodologies)
 - Twin dilemmas -- information overload *and* gaps
- **HIU initiative: Visualized Information and Synthesized Temporal Analysis (VISTA)**



Selected Non-DoD Insights

- **NOAA advice**

- “Get librarians involved”
- “Metadata are hard -- start now!”
- “If the metadata are not good, the data might as well be lost”
- “Use non-proprietary tools”

- **NSF observations**

- Share knowledge via partial models
- Policy and technology are partners in research
- There is a potential need for a new data model beyond relational to support graph bases, queries

- **NIH insight**

- “Researchers are as willing to use other people’s data models as they are to wear other people’s underwear”



Additional Insights From the Plenary (1 of 2)

- The data problems that the DoD are facing are *not* unique to the DoD
 - many other government agencies seem to have the same issues, including, *inter alia*,
 - Data sharing
 - Data purity
 - Metadata policy (e.g., standardization)
 - Data shelf life
 - Data naming conventions
 - Data reconciliation
 - Data maintenance
 - Ontological development for intelligent searches
 - Data protection
 - Data provenance
 - Data surrogation
 - Data bloat



Additional Insights from the Plenary (2 of 2)

- The primary issues that the DoD must address to achieve the strategic vision are *policy, engineering, social, and and informational* (vice research and development); e.g.,
 - For both traditional and non-traditional partners, agreement on
 - Need-to-know
 - Standards
 - Security conventions
 - Development of naming conventions, ontologies, common semantics and syntax



Agenda

- **Panel Goals, Composition**
- **Insights on Nature of the Problem**
- **Preliminary Findings, Recommendations**
- **Summary**

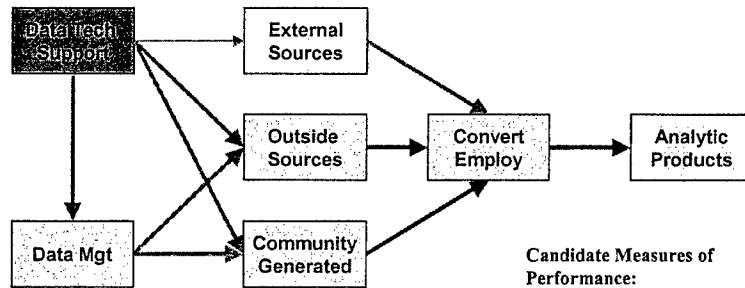


This section describes the preliminary findings and recommendations that have emerged from the workshop.



The Working Groups: An Alternate View

What limits this network?



Candidate Measures of Performance:

Timeliness
Availability
Usefulness
Transparent
Risk Assessed



This diagram was designed to show supporting dependencies among the working groups.

WG 1: Data Management

- **Perceived nature of the problem**
 - Data Management Issue is vast, uneven in maturity across the DoD, and lacks a comprehensive working process to meet the performance measures identified by the Plenary Speakers
- **Best practices and recommendations**
 - Move from a centralized to a more decentralized (networked) repository structure
 - Need to define a disciplined framework (a la JSF example)
 - Recommend committee to establish and manage framework
 - Improve the business model surrounding data
 - Existing situation lacks clear incentives, has disincentives
 - Time to maintain and productize, cost, negative impacts of bad outcomes
 - Incentivize (financial, report card, measures, considerations)
 - Establish process and procedure for re-use and risk management
 - Extend policy and enforce use of Metadata



WG 2: Obtaining Data from Outside Sources

- **Perceived Nature of the Problem**

- Definition of data needs comes from the mission-task statements as a function of level of “war.” Qualified by distinction between normal operations (peace) and response operations (crises).

- **Best Practices and Recommendations**

- Relate data applications (core, common, custom) to users (providers-sources) (enterprise DoD, communities of interest, private)

- **Further Issues to Pursue**

- Mission – task statements roll down (strategic, operational, tactical);
Can data sets be rolled up, once defined at the lowest tactical level?
- Need clarity on two triads:
 - “Core, common, custom”
 - Enterprise, Communities of Interest (Col), Private



Working Group 3

- **Major obstacles/barriers**

- State of art of performance modeling (individual; organizational)
- Lack of unified theory of war
- Data aggregation - nonlinear effects
- Large community of data providers – presents a large info management problem



Working Group 3 (Concluded)

- **Some promising best practices**
 - AMSAA & JDS concepts
 - Data audit trail/pedigree process
- **Follow on issues for MORS**
 - Details of data standards
 - Nomenclature/operational definitions/categories
 - Resolution/fidelity
 - Format/tagging/metadata
 - Linkage of modeling paradigms and data standards
 - Process for data requests



Working Group 4

- **Perceived nature of the problem**
 - Internal data management and generation
 - Trust issues with sharing (misuse and risk to provider)
- **Best practices/recommendations**
 - Organize within provider organizations; don't standardize among organizations
 - Provide external structure for sharing
- **Further issues to pursue**
 - Aggressively work the trust issue; DODD/DODI and analytical baselines is the place to start

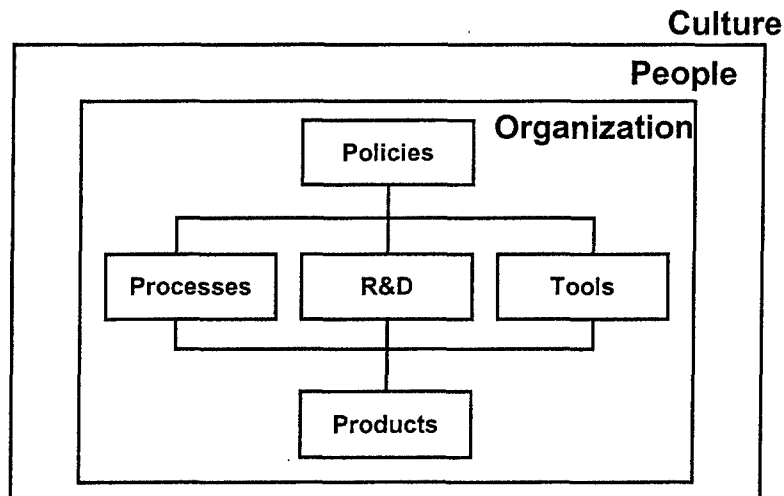


Working Group 5

- **Perceived nature of the problem**
 - Understand proper relationships among analytical processes (e.g., survivability analysis), data business processes (e.g., create threat data), data technical processes (e.g., rapid generation of data from engineering models), and data tools (e.g., MATLAB/SimuLINK)
- **Recommendations**
 - General: Apply standard system development theory (with first step to identify business processes to be supported) in the use of existing technologies or development of new technologies
 - Specific: Standardize generation of metadata
- **Further issues to pursue**
 - Further define business processes
 - Consider input from other working groups on any recommended data technical processes or tools



Framework for Recommendations



The accompanying figure provides a business process re-engineering perspective of the data-analysis enterprise. The Synthesis Panel concluded that if enhanced data practices are to improve defense analysis, we must consistently address *all* of these factors. The backdrop for these factors is set by the *cultures* of the many communities that must participate in the practice and use of data for analysis. It was recognized that in many cases DoD would rely on the data provided by non-traditional partners (e.g. NGOs). Thus, we must be cognizant of the cultures of the other participants and flexible in our interactions with them.

Second, *people* are critical components of the data enterprise. This implies the need to provide critical Education and Training (E&T) for both the providers and users of the data. Third, we may have to change *organizational* relationships to facilitate the sharing of information among members of the community.

Within this framework, key *policies* must be formulated that seek to overcome residual barriers to the sharing of key data. DoD Directive 8260 constitutes an important initial step, but more needs to be done.

Finally, key *products* are needed to explore and refine our understanding of key new data concepts (e.g. the analytical baseline).

Key Recommendations (1 of 4)

- **Culture**

- A *fundamental* change in the data culture is required (e.g., power is derived from *sharing* vice *hoarding* data)
- Accelerate actions (e.g., meetings, coordination efforts, socialization) to *breakdown barriers with the diverse communities* who must participate in the data enterprise



Key recommendations, or observations, as viewed by the Synthesis Panel are broken out using the preceding framework. The notion that power is derived from sharing data — rather than simply that “knowledge is power” — necessitates a culture change.

Key Recommendations (2 of 4)

- **People -- Analysts**

- Develop curricula, programs to *enhance education and training for the military operations analyst*, emphasizing the criticality of data in the analysis process

- **People -- Decision makers**

- Institutionalize the commitment of senior decision makers to address the data problem
- Provide decision makers with a list of data-related questions that they should pose to the analyst team (see next slide for strawman)



The Education and Training (E&T) of providers and users of data is viewed as a critical enabler.

Strawman Questions for the Decision Maker to Pose to the Analyst

- **Prior to the Study**
 - Do you need any data/information (or authorization) from me?
- **At the Initial Review**
 - What do you perceive the "real" issues to be?
 - What data will you need to illuminate the "real" issues? Have you generated and coordinated a plan to address the data issue?
 - What barriers (if any) do you perceive in order to access, collect, generate, or transform the data required? How does your data plan address these barriers?
- **At the First Iteration**
 - Are you adhering to your data plan?
 - If not, why not?
- **At Final Report**
 - What key lessons did you learn with respect to data?
 - What steps are you taking to make your data *accessible and useful* to the greater community?



This is a strawman set of questions that the analyst needs to be prepared to answer at each milestone of an analysis.

Key Recommendations (3 of 4)

- **Organization**

- Establish organizational mechanisms to encourage Interagency, International cooperation on data sharing

- **Policies**

- Reassess existing policies which severely restrict the flow of data, information across institutional barriers -- *rebalancing* security concerns and the “need-to-know” [should we re-examine the existing “need-to-know” policy in which there is a presumption of guilt, vice innocence?]

- **Tools**

- Expand the analyst’s “tool chest” to support the collection, generation, transformation, V&V, and visualization of data



If defense analysis is to be improved through better data practices, enhanced interagency and international cooperation is essential.

Policies are needed that rebalance security concerns and the “need-to-know.”

Tools that help to rapidly synthesize and display the most critical data to the decision maker are needed.

Key Recommendations (4 of 4)

- **Processes**

- Develop a data support business process that exploits strengths (e.g., encourages the generation of metadata), ameliorates weaknesses (deals with disincentives such as proprietary concerns)
- Explore options to facilitate the tasking of the Intelligence Community to acquire/provide needed data
- Convene a NATO Studies, Analysis and Simulations (SAS) Panel to develop an alliance CoBP on data for analysis (analogous to C2 Assessment and OOTW CoBPs)

- **Products**

- Perform pilot studies to clarify the desired attributes of the Analytical Baselines
- Continue to establish repositories, data warehouses to archive, provide access to V&V'ed data, for those with a validated need



Data support business practices need to reinforce key strengths and ameliorate potential weaknesses. Organizations first need to get their internal processes in order than focus on external sharing of data.

Desired attributes of analytical baselines need to be tested to ensure they support the full range of assessments. Repositories may need to be expanded to meet data needs at a broader range of assessment activities (e.g. simulation based acquisition).

Thoughts on Best Practices: Data

- **Employ the Data Taxonomy identified in the NATO CoBP for C2 Assessment**
- **The analysis team must determine**
 - What data are needed in which structure
 - Who owns these data
 - Security issues
 - Costs to buy, collect, or generate data
- **In the absence of needed data, it is good practice to use the knowledge of subject matter experts to generate the needed data**



Thoughts on Best Practices: Metadata, Common Data Infrastructure

- **Metadata -- “information about information”**
 - The source of the data, the reliability, and assorted assumptions and constraints *must* be captured in standardized metadata sets
- **Common Data Infrastructure --**
 - Adhere to data engineering principles to contribute to data reuse
 - Archive data in retrievable form using standardized meta-data sets



Agenda

- **Panel Goals, Composition**
- **Insights on Nature of the Problem**
- **Key Findings, Recommendations**
- **Summary**



This section briefly summarizes some of the Synthesis Panel's major observations and conclusions.

Summary

- The data problem is complex and enormous ... and increasing in both complexity and size!
- The community has taken significant *initial* steps to address the problem; e.g.,
 - Issued new directives, instructions (e.g., 8260)
 - Created new organizations (e.g., JADMSC)
 - Promulgated new tools, standards (e.g., DDMS)
- However, in order to make further substantive improvements, we have to
 - Transform the *culture* (e.g., by implementing incentives, overcoming disincentives)
 - *Educate and train* the users, providers of data
 - Implement new *processes* (e.g., work the metadata problem)



Basically, the data challenge is a huge one, and, while the community has taken several significant steps to address it, more are needed – as illustrated by the actions listed here: cultural transformation, E&T, and new processes to solve the data problems.

Scorecard

- **Accomplishments:**

- The Workshop served to clarify the nature of the data problem
- To paraphrase the psychiatrist in "Portnoy's Complaint": "Now, we are ready to begin..."

- **Residual Challenges**

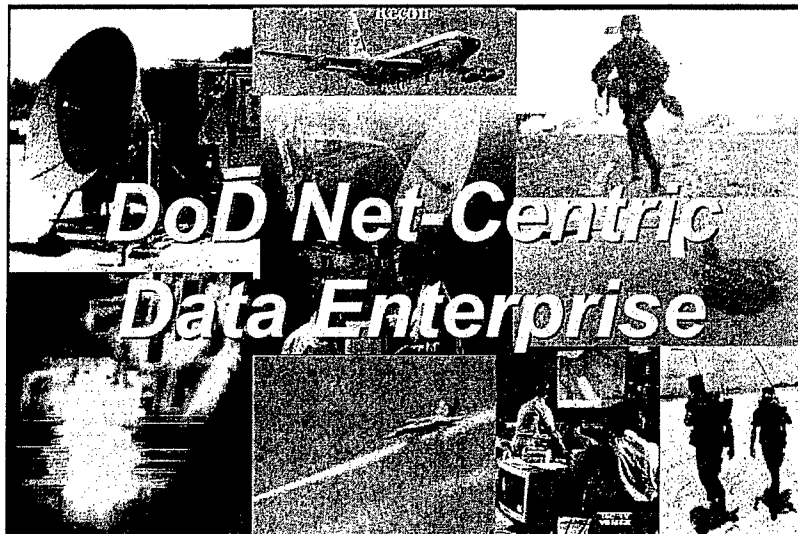
- There is a need to reach out to a broader community (e.g., inter-agency, coalition) to address the data problem
- There is a need to generate a plan of action and milestones (POA&M) to focus, mobilize the community
- Regard this workshop as the first in a sequence



The Synthesis Panel sought to put the workshop into perspective by generating a scorecard to summarize the positive accomplishments of the event and the residual challenges that confront the community.

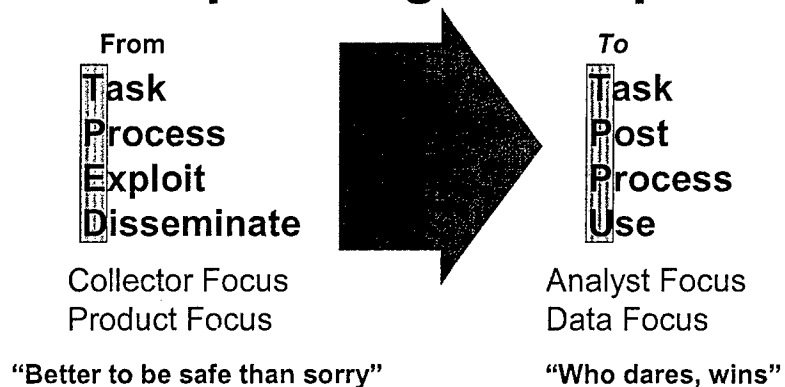
The workshop made good-to-excellent progress on our four objectives. However, it must be understood that this constitutes a beginning, not an end. Our status is reminiscent of Philip Roth's Portnoy at the conclusion of his book "Portnoy's Complaint". After a difficult, soul-wrenching confessional, his psychiatrist concludes the book with the words "Now, we are ready to begin..."

Since we are now "ready to begin", we believe that there are three major initiatives that the community should undertake. First, we need to reach out to a broader community. Second, we need to generate a POA&M to focus and mobilize the analysis-data community. Finally, MORS should plan follow-on workshops to address the critical issues identified during this meeting.

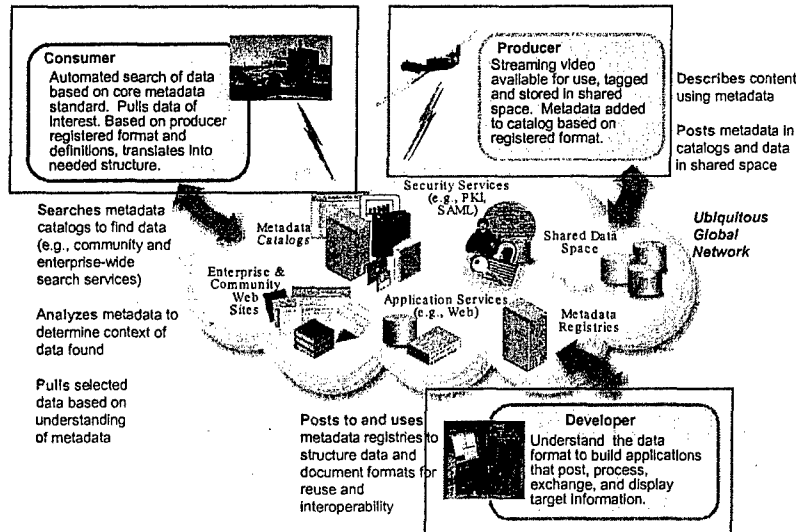


John.osterholz@osd.mil

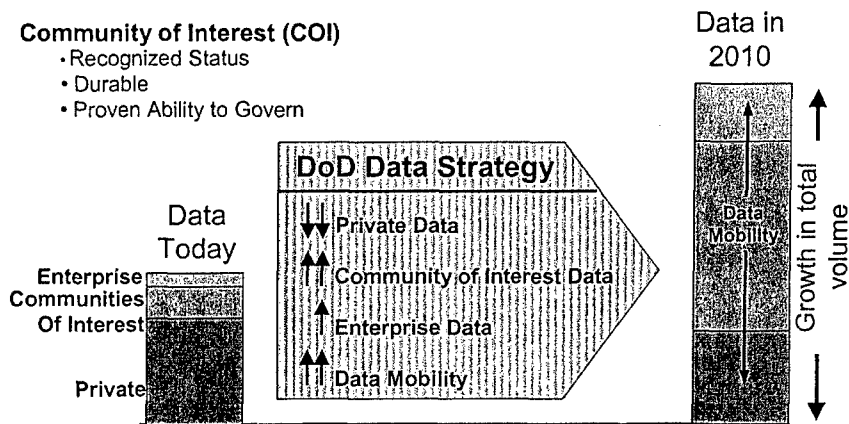
The New Enterprise Data Operating Concept



Making Data Accessible



Enterprise Data Strategy



"Interoperability begins at the data level, not the systems level, have a common language, common tagging. Data is our most important product."

Brig. Gen. Mike Ennis,
Marine Corps director of intelligence,

Enterprise Data Strategy Goals

Department of Defense
Office of Deputy Chief
Information Management
Directorate

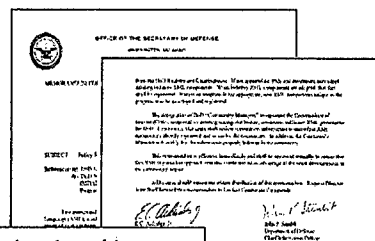
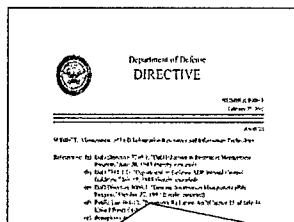
Enterprise Strategy for Data
Management



February 18, 2003

Goal	Description
Goals to increase percentage of Enterprise and Community data over private user, system and application data	
Visible	Users and applications can discover the existence of data assets through catalogs, registries, and other search services. All data assets (intelligence, non-intelligence, raw, and processed) are advertised or "made visible" by providing metadata, which describes the asset.
Accessible	Users and applications post data to a "shared space". Posting data implies that a) descriptive information about the asset (metadata) has been provided to a catalog that is visible to the enterprise and b) the data is stored such that users and applications in the enterprise can access it. Data assets are made available to any user or application, except when limited by policy, regulation or security.
Institutionalized	Data management approaches are incorporated into Department processes and practices. Metrics and incentives are established to assess progress against data management strategy implementation and the success of the data management approaches.
Goals to increase use of Enterprise and Community data	
Understandable	Users and applications can comprehend the data, both structurally and semantically, and readily determine how the data may be used for their specific needs.
Trusted	Users and applications can determine and assess the authority of the source because the pedigree and security level of each data asset is known and available.
Interoperable	Facilitation of many-to-many exchanges of data between pre-defined or unanticipated systems. Hence, sufficient information about the data must be known to allow mediation or translation of data between these interfaces.
Responsive to User Needs	User perspectives are incorporated into data management approaches via feedback to ensure satisfaction.

DoD Direction on Data



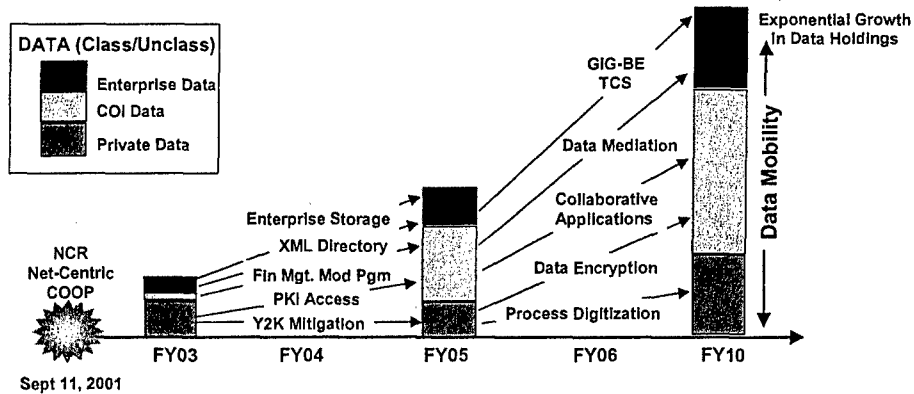
4.4.2. Data and information shall be structured to enable full interoperability and integration across DoD operations and activities. Creation of duplicate data shall be minimized and to the extent possible data shall be entered only once. This increases data accuracy and decreases several sources of data.

The DoD XML Registry and Clearinghouse is the authoritative source for XML components. All Program Managers that use XML as an interchange form must register their XML components in accordance with procedures established by DISA.

DoD Enterprise Data Strategy Implementation

"Interoperability begins at the data level, not the systems level, have a common language, common tagging. Data is our most important product."

Brig. Gen. Mike Ennis,
Marine Corps director of intelligence,





Analytic Agenda

Dr. Christopher Lamb

March 25, 2003

DASD(Resources and Plans)



Outline

- **Larger context**
- **Strategy and its elements**
- **Roles**
 - OSD(Policy)
 - OSD(PA&E)
 - Joint Staff J-8
- **Issues**

DASD(Resources and Plans)



Larger Context

- In order for transformation to succeed, certain legacy methods of doing business must be reformed. The areas most in need of reform are:
 - **Acquisition:** The *time* it takes to acquire major systems must be sharply reduced, without significantly increasing technical, financial, or programmatic risk
 - **Requirements:** The CJCS must be given the analytic resources that will enable him to determine *priorities* among competing military requirements
 - **Trades:** The Secretary must be provided with the integrated multi-disciplinary information necessary to make big *trades* across large sectors of defense, allowing him to explicitly balance costs and risks
- A prerequisite to achieving the Trades goal is transforming the analytic system.

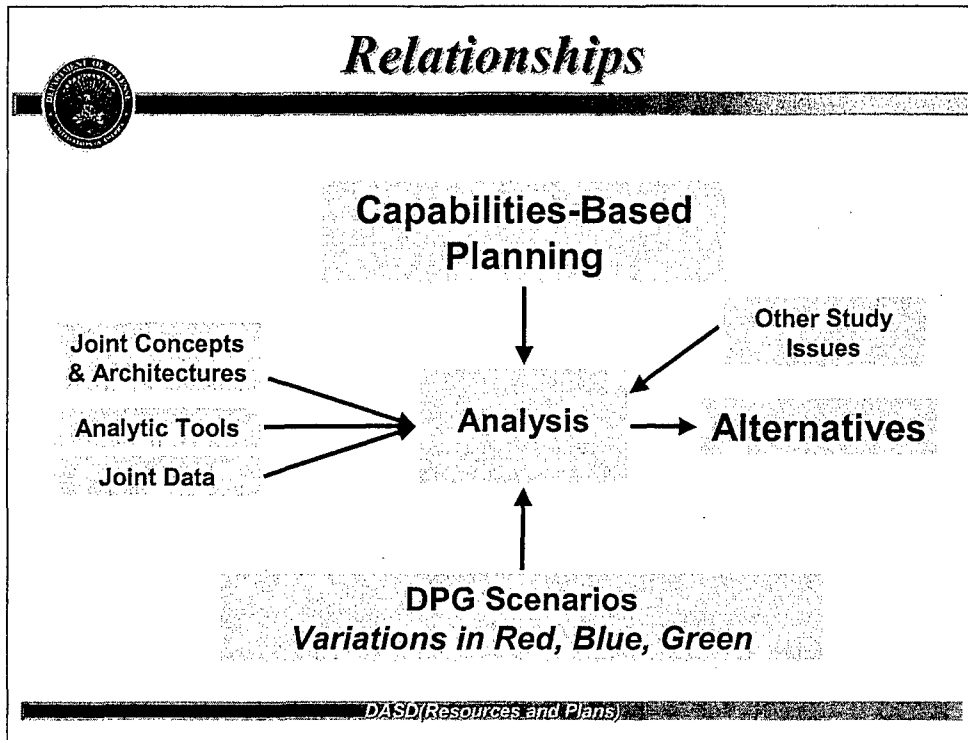
DASD(Resources and Plans)



Transforming the Analytic System

- **What is needed**
 - Higher quality, better synchronized, more anticipatory/proactive, more responsive, more flexible, more cooperative, more efficiently conducted analytic support for senior leader decision making
 - Keyed to their current interests and concerns
 - Effective response to unanticipated needs
 - Analytic system that facilitates transformation of the Department and successful implementation of the new Defense Strategy
 - All the analytic support needed to help ensure that our forces are robust in the face of an uncertain future and enjoy Full Spectrum Dominance
- **How we get it: the Analytic Agenda- an agreement and plan between OSD Policy, OSD PA&E, JS to accomplish this**
 - Capabilities-Based Planning
 - Joint Data
 - Joint Operational Concepts
 - Analytic Resources
 - Wider array of approved tools

DASD(Resources and Plans)



**Aligning Analysis with
Capabilities-Based Planning**

- **What is needed:** Ability to explore a wider range of variation in planning to better account for future uncertainties
 - Analyses that help “build a portfolio of capabilities that is robust across the spectrum of possible force requirements, both functional and geographical.” (QDR 2001)
 - Variation in threat, Blue/Green capabilities, force size, composition, technological sophistication
- **How we get it:** A uniform, Department-wide approach to Capabilities-Based Planning that accounts much better for uncertainty (issued by Policy)
 - Provides a wider range of variation than traditional approaches
 - Analytically tractable with reasonable demands on available analytic resources
 - Politically defensible with Congress in terms of plausibility
 - Also need to transform DPG scenarios to facilitate Capabilities-Based Planning
 - Size the force using DPG scenarios, varying all key parameters through their full range of uncertainty

DASD(Resources and Plans)



Joint Operating Concepts

- What is needed

- Uniform, integrated, agreed-upon descriptions of how the transformed military will conduct joint operations in the future, for representation in analytic efforts
 - Warfighting
 - Other operations

- How we get it

- Joint Staff with JFCOM to complete current effort to establish JOCs and related hierarchical operational constructs

DASD(Resources and Plans)



Tools and Methodologies

- What is needed

- A broader range of analytic tools and methodologies to enable strategic decision making by top leadership and better suited to Capabilities-Based Planning
- Better analytic methods for Combatant Commanders
 - Support war planning and readiness/risk assessments, development of alternative Courses of Action, exploration of various settings of risk management "dials"
- Authentic analytic coverage of prominent areas like WMD, IO, C4ISR, TMD

- How we get it: establish broader set of tools/methodologies

- JAMIP identify specific shortfalls, oversee development
- Tools and study methodologies explicitly accounting for the new Risk framework
- Allows "aggregate-up" analysis plus traditional micro-level analysis
- Oriented toward new DPG Scenario structures when established

DASD(Resources and Plans)



Data

- **What is needed:** Assurance that data needed for strategic analyses will be available, transparent, accurate, standardized, sufficiently detailed, and synchronized with the PPBS and QDR cycles
- **How we get it**
 - Scenario data management Directive and Instruction (both have been approved) institutionalizing an aggressive, systematic, dynamically managed approach ensuring:
 - On-the-shelf availability and currency of baseline data sets
 - Lock-step synchronization and on-time delivery of major contributors to analytic data sets
 - Ensure that at least the following data are available when needed for analyses:
 - Near-term: Data needed to assess executability of war plans
 - End of FYDP: Projected Blue/Green forces data, threat projections
 - Longer range: Blue and Red data covering 10 years post-FYDP

DASD(Resources and Plans)



Resources

- **What is needed:** Sufficient skilled manpower, funding, and facilities to conduct key recurring and one-time activities in Washington and by Combatant Commanders
 - OSD, Joint Staff, Services, COCOMs
 - Intelligence Community
- **How we get it**
 - OSD + Joint Staff + DIA + Services to collaborate more closely
 - Near-term: PDM to rectify shortfalls in analytic resources (COCOMS + OSD/JS)
 - Longer term: Institutionalize multi-year plan to support Analytic Agenda and POM to implement that plan

DASD(Resources and Plans)



The Players

- OSD (Policy)
- OSD (PA&E)
- Joint Staff
- The Services
- Combatant Commanders

DASD(Resources and Plans)



OSD (Policy)

- **Developing a new approach to supplant the existing Defense Planning Guidance Illustrative Planning Scenarios**
 - Provide for much greater variability than traditional excursion approach
 - Goes together with Capabilities-Based Planning, approach also under development
- **Briefing series for senior OSD leadership: Big Picture, Analytic Agenda, Scenarios, Major Policy Issues**
 - Recommend new scenario structure and major components
 - Planning cases -- what scenarios we should have (e.g., Major Combat Operations, HLD, SSCs, Far Term)
 - One fully developed scenario as proof-of-concept
- **After approval by Department leadership**
 - Develop and coordinate Terms of Reference for actual scenarios
 - Build and coordinate actual scenarios
 - Present scenarios for Secretary approval this Spring

DASD(Resources and Plans)



OSD(PA&E): Data Directive & Instruction

- DoD Directive: "Data Collection, Development, and Management in Support of Strategic Analysis"
(was approved in December)
 - OSD Policy: Identify and prioritize scenarios
 - Joint Staff: Lead development of current year analytical baselines
 - OSD PA&E
 - Lead development of future year analytical baselines
 - Manage current and future year data repository (*JDS*)
 - DoD Components
 - Provide data
 - Participate in development of analytical baselines
 - Creates the Joint Analytic Data Management Steering Committee (JADM-SC)
- DoD Instruction (same title) approved in January:
Detailed implementation instructions for the Directive

DASD(Resources and Plans)



OSD(PA&E): Analytical Baselines

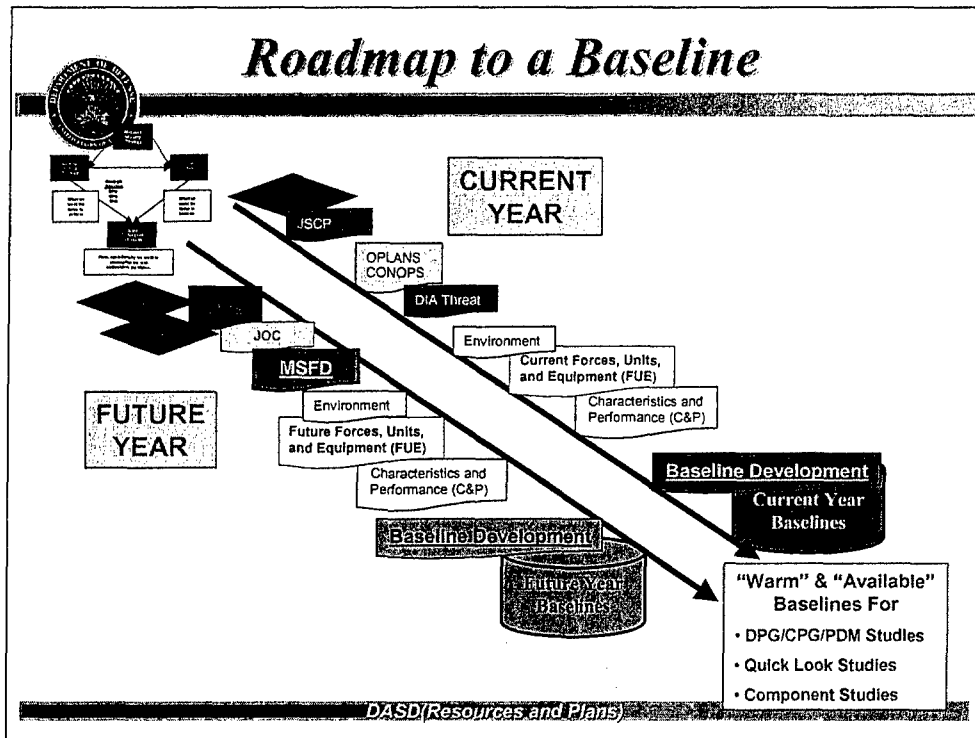
Definition: A package comprised of a scenario, concept of operations, and integrated data used by the DoD Components as a foundation for "strategic analyses"

Scenario	CONOPS	Other Data
•Location	•Red	•Characteristics and Performance
•Year	•Blue	•Forces, Units, and Equipment (FUE)
•Objectives	•Green	•Environmental (terrain, weather, ...)

- Includes analysis
- If models are used in the analysis the Baseline will include model input data, the model (if available), and output results

Goal: Visible, accessible and traceable data ready to support Department-level analyses

DASD(Resources and Plans)



OSD (PA&E) Other Support to Analytic Agenda

- Support Joint Staff Operational Availability Study
- Take the lead on Transformation Force Assessment (end-of-FYDP force), working with Joint Staff
- Division of labor with Joint Staff
 - Joint Staff focuses efforts on Near-Term
 - PA&E focuses efforts on Mid/Long-Term

DASD(Resources and Plans)



Joint Staff J-8

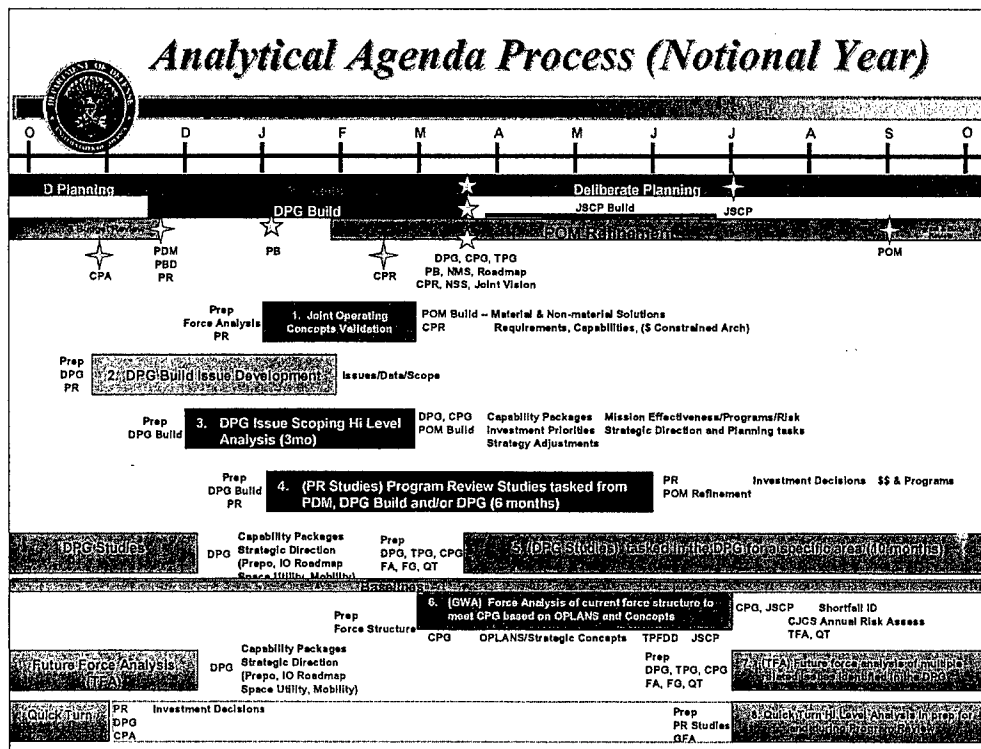
Annual Warfighting Assessment Program

- **Definition:** A combination of quantitative and qualitative analysis conducted by JS and OSD to assess current and future warfighting capabilities
- **Purpose:** Provide DoD leadership with properly focused and sequenced analyses to help inform strategic and programmatic decision-making processes

Major Continuing/Periodic Activities

- **Global Warfighting Assessment:** Assess capability of current forces to execute CPG
- **Transformation Force Assessment:** In support of OSD PA&E, assess end-of-POM force as input to DPG and TPG
- **Quick Look:** Short turn analyses as needed by senior leaders

DASD(Resources and Plans)





Issues

DASD(Resources and Plans)

Unclassified



Model Adequacy Assessment

• ISSUE

- The transformation of Defense analysis that we must undertake calls into question the adequacy of existing and developmental models and methodologies
- The degree to which they will be able to assess two key classes of analytic questions must be assessed
 - Higher level, more highly aggregated, complex, multi-disciplinary questions
 - More subtle, difficult-to-model phenomena of increasing importance, e.g., Special Operations, C4ISR, IO, WMD effects, Urban Operations

• RECOMMENDATION

- Comprehensively identify the types of issues, phenomena, and analyses that we expect to be required to support
- Systematically compare those needs with the capabilities and limitations of existing and emerging models (e.g., JWARS) and methodologies
- Determine steps that must be undertaken to rectify those shortfalls
- Establish programs and funding to implement those steps

DASD(Resources and Plans)

Unclassified



Related Issue: Multi-Resolution Modeling

• ISSUE

- The more complex, multi-disciplinary questions being addressed by Department leadership raise questions about the resolution level of analytic models
- A layered, multi-resolution approach is probably called for:
 - Broader, lower-resolution models first to help scope out complex issues
 - Narrower, higher-resolution models next to provide deeper insights
- The adequacy of available scoping models is questionable

• RECOMMENDATION

- Include in the assessment described on the previous slide explicit evaluation of the needs for multi-resolution modeling

DASD(Resources and Plans)

Unclassified



Comparability of Study Results

• ISSUE

- Strategic analyses of similar/overlapping issues often yield incompatible results
- Top leadership decision making on major issues requires an understanding of the reasons for this
- Current approach to evaluating these reasons is piecemeal and issue-by-issue

• RECOMMENDATION

- Evaluate several recent studies of overlapping substantive issues to identify the strategic causes of incompatible results
- Develop and institutionalize a more systematic means of capturing the assumptions, parameter settings, and methodological approaches used in strategic analyses, in a way that facilitates comparability analysis across studies
- This is envisioned as an analytic framework *above* the level of Analytic Baselines, and capturing their relevant contents horizontally

DASD(Resources and Plans)

Unclassified

UNCLASSIFIED



Baselines and the Analytic Agenda

Dr. Jim Stevens, Director, Joint Data Support, OSD PA&E

March 27, 2003

UNCLASSIFIED

UNCLASSIFIED



What's an Analytical Baseline?

- Warm databases ready to support Department-level analyses
 - CPG or DPG based data for studies and analyses
 - Developed by a collaborative team led by the Joint Staff (current-year) or PA&E (future-year)
- A package comprising a scenario, concept of operations, and integrated data used by the DoD Components as a foundation for strategic analyses (DoD Directive 8260.1)
- For example, OA Study baselines will be developed for both the SWA and Korea 2010 SDTE scenarios with

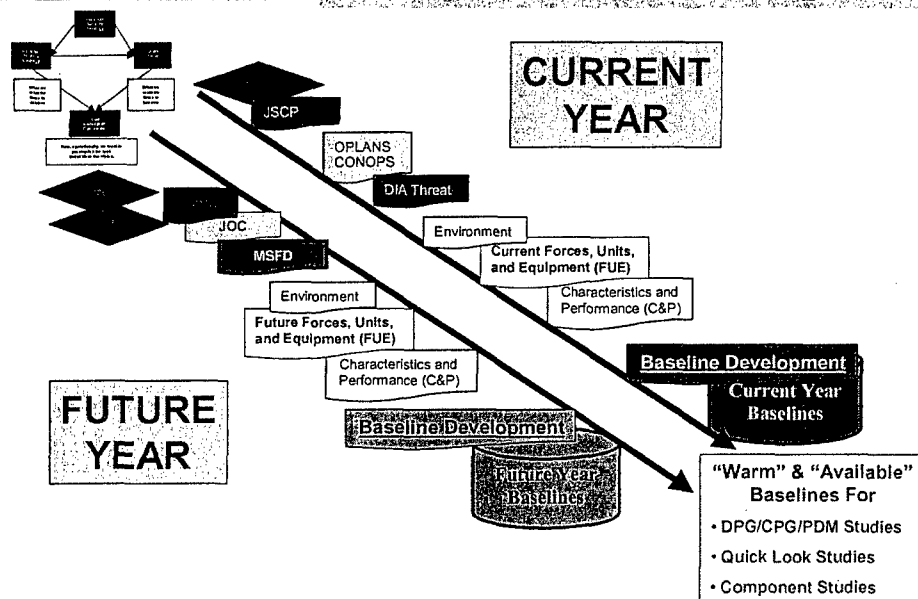
Scenario	CONOPS	Other Data
•Location	•Red	•Characteristics and Performance (C&P)
•Year	•Blue	•Forces, Units, and Equipment (FUE)
•Objectives	•Green	•Environmental (terrain, weather, ...)

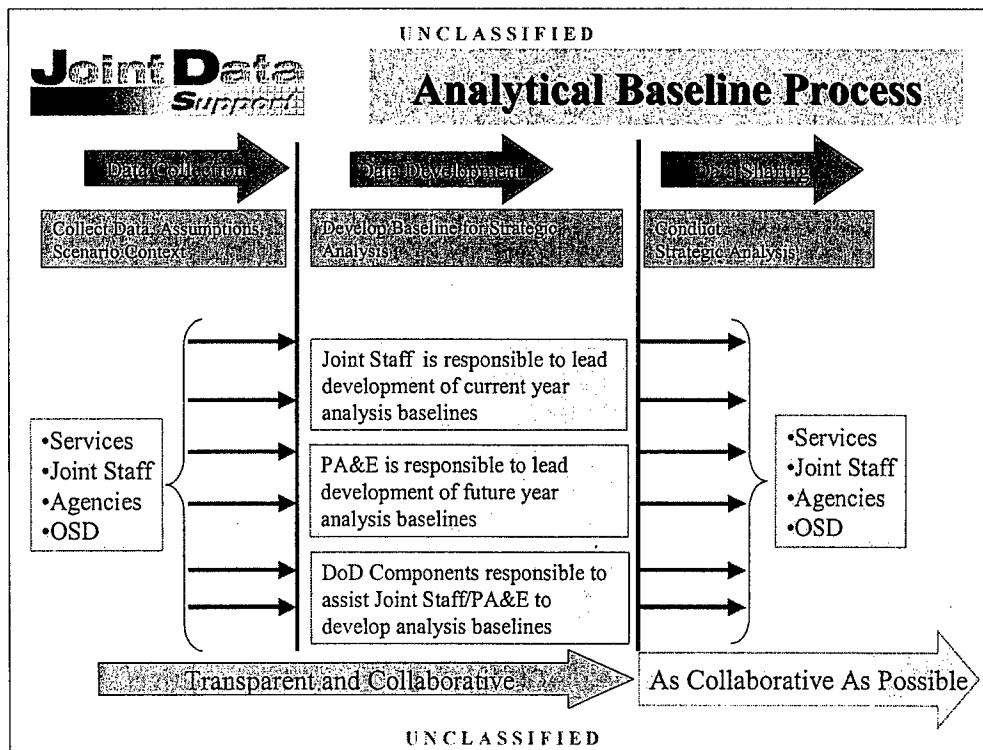
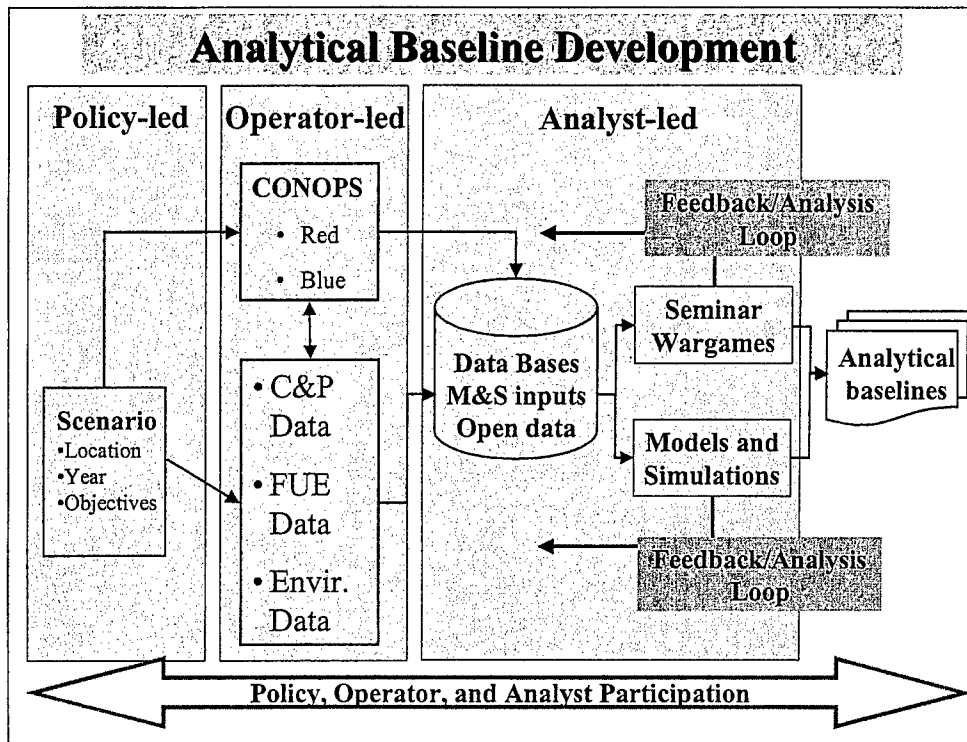
UNCLASSIFIED

Who Uses a Baseline?

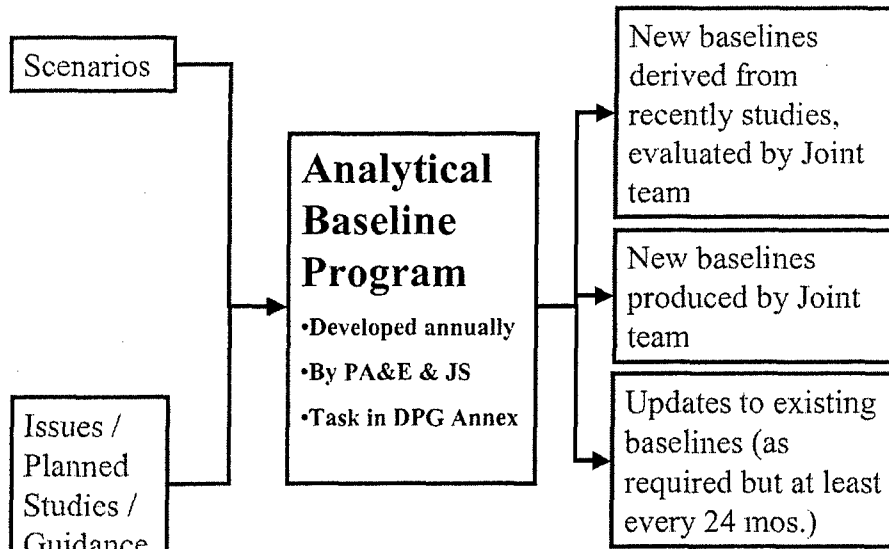
- Should be used by DoD Components as a departure point for analysis
 - Campaign studies and analysis
 - Department-level study excursions and refinement
 - Component-focused “how we fight” studies
 - Acquisition programs: Examples
 - GMTI
 - F22
 - JSF
 - Minuteman
 - Operational Test & Evaluation

Baseline Roadmap

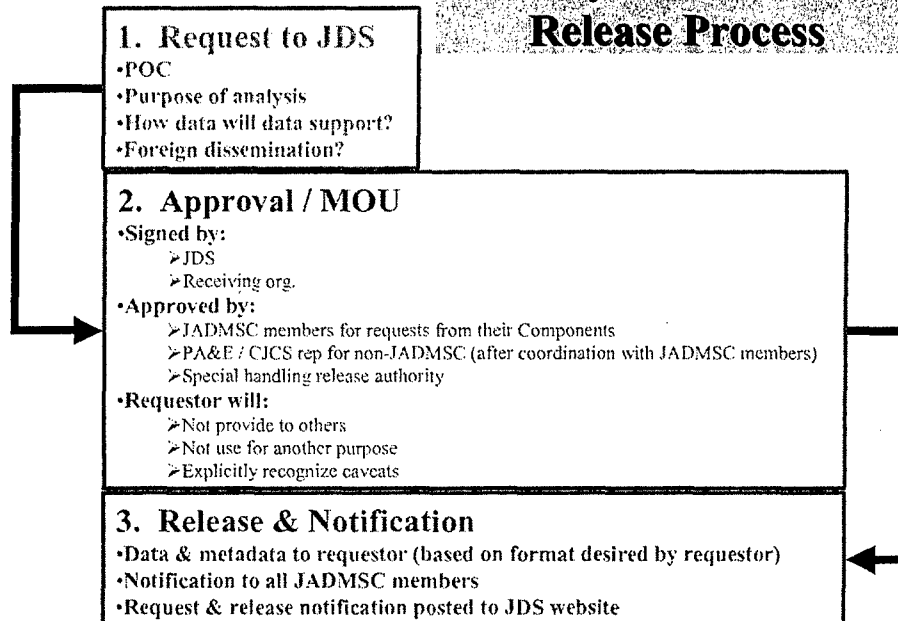




Develop/Maintain A Baseline



Analytical Baseline Release Process





UNCLASSIFIED

2010 Baseline (Based on the OA Study)

Proposed Development Schedule

- **March**
 - Identify Baseline development team lead
 - Asks JADM SC to appoint team members
- **March-April**
 - OA Study complete
 - Initial team meeting
 - Baseline preparation
- **May**
 - Baseline provided to JADM SC for review and comment
 - Analytical baseline updated and JADM SC briefed
- **Late May**
 - JDS publishes 2010 Analytical Baseline

Proposed Table of Contents

- General description of the scenario
- Results for the baseline run
- Assumptions, caveats, and limitations
- Databases
- Models descriptions/versions

UNCLASSIFIED



UNCLASSIFIED

Role of Joint Data Support (JDS)

- **Helps identify, collect, and develop data supporting analytical baselines**
 - Specifies data requirements
 - Issues data calls
 - Performs quality checks (V&V)
 - Archives and maintains pedigree for all data
- **Manage and disseminate analytical baselines**
 - Member of all Analytical Baseline development teams
 - Single distribution point for all baselines
 - Helps ensure appropriate use of baseline data
- **Assist in the administration of ongoing studies**
- **Repository of completed studies, models, and analyses**

UNCLASSIFIED

Backup slides

Summary – Value to the Department

Better analysis that provides a deeper understanding of the key issues

- **A common starting point can result in**
 - Quicker response to questions
 - Better understanding of differences among competing analyses
 - More time for analysis as less time is needed for data collection
- **Better data via collaborative development of Baselines**
 - Data from best available sources (to include meta data)
 - Extensive review reduces database errors: greater consistency and jointness
 - Economy of resources via pooled efforts to develop the “tough” data

Visible, accessible and traceable data ready to support analysis



UNCLASSIFIED

The Use of Analytical Baselines

- **Components are encouraged to use data from analytical baselines in their analyses**
- **When using an model-based analytical baseline the user should**
 - Run the model with baseline data to verify model parameters and model version
 - Explain differences between their study results and those provided with the analytical baselines
- **Studies using analytical baselines will be asked to identify what changes they made to the data in the baseline during the conduct of their analysis**
 - Permits focus on assumptions vice models and overall database
 - Basis for comparing analyses

UNCLASSIFIED



UNCLASSIFIED

JDS Charter: Improve the Quality and Consistency of DoD Analyses

- **Support the Analytic Agenda - Identify, collect, develop, manage, disseminate data and associated analytical baselines**
 - Using JAMIP simulations (e.g., ongoing DPG-directed Operational Availability Study)
 - Using other decision support processes (e.g., Dynamic Commitment, QDR)
- **Support Joint Warfare System (JWARS) Development and Fielding**

UNCLASSIFIED



UNCLASSIFIED

Analytic Agenda Rationale

- **Previous analysis process**
 - Focused on a 2 MTW strategy
 - Analysis was 'event-based', focused around major multi-year studies
 - Each new study re-invented basic data and trained/prepared analysts
 - Forces-based versus capabilities-based deployment and conflict
 - Limited use and re-use
- **The new 1/4/2/1 strategy requires greater analytic flexibility**
 - More scenarios, larger number of permutations = greater workload
 - Scope of the analytic problem similar to complexity of OPLAN deliberate planning goals
- **The Department's analysis goals**
 - Responsive
 - High quality
 - Capable of training and developing analysts (rotating military)
 - Imaginative in developing new ways to assess the strategy and force

UNCLASSIFIED



UNCLASSIFIED

New Method Must Provide

- **A quick start to new studies**
- **An open process that allows discussion and review**
 - Consensus not required
 - Understand differences
- **A departure point for constant development and refinement**
- **Better analysis that provides a deeper understanding of the key issues**
- **More time for analysis and less time for data collection**
- **The ability to develop, understand, and study Joint Operating Concepts**

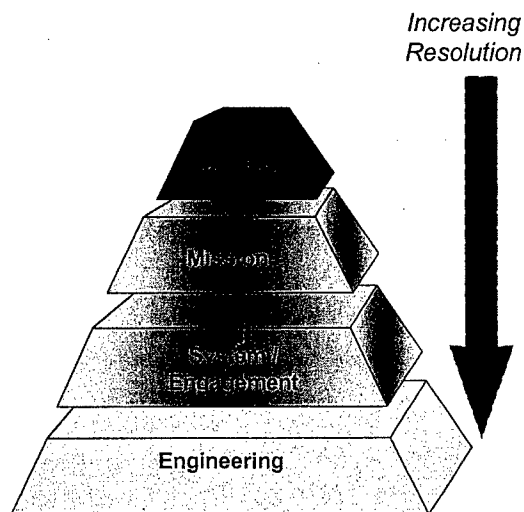
UNCLASSIFIED

Joint Future CONOPS

- Starting point normally USD Policy developed scenarios that include strategic objectives and a strategic concept
- Maximize operator involvement
 - Joint Staff (lead)
 - Combatant Commanders
 - Services

UNCLASSIFIED

Modeling & Simulation Paradigm



UNCLASSIFIED

Objective

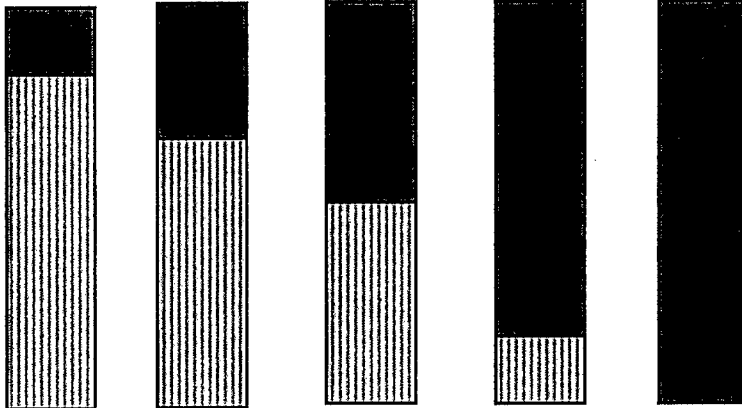
- **Create databases in a collaborative manner that are available to those that need them**
 - Reduce duplication in database development efforts
 - Greater focus on identifying and resolving database deficiencies
 - Encourage use of common databases
- **Provide “warm” model ready databases, when appropriate**
 - Help resolve inconsistencies among data providers
 - Allows quick start for analysis

Definitions

- **Concept of Operations (CONOPS):** A verbal or graphic statement, in broad outline, of a commander's assumptions or intent in regard to an operation or series of operations. The CONOPS is frequently embodied in campaign plans and operation plans, in the latter case particularly when the plans cover a series of connected operations to be carried out simultaneously or in succession....
- **Force:** An aggregation of military personnel, weapon systems, equipment, and necessary support, or a combination thereof.
- **Unit:** Any military element whose structure is prescribed by competent authority, such as a table of organization and equipment; specifically, part of an organization.
- **Equipment:** In logistics, all nonexpendable items needed to outfit or equip an individual or organization.
- **Strategic Analysis:** An analysis of force sufficiency and effectiveness conducted by DoD Components to support the development and evaluation of the defense strategy. Such analyses address both forces and enablers (e.g., inter-theater and intra-theater lift capability).

Notional Data Categories Available

- | | | | | |
|--|---|---|---|---|
| FUE <ul style="list-style-type: none"> • Forces • Units • Equipment • Munitions • Platforms • C&P | Performance <ul style="list-style-type: none"> • Adjudication • Network • Performance • Training Level • Infrastructure • Sortie Rates | Environment <ul style="list-style-type: none"> • Terrain • Weather • Oceanography | CONOPS <ul style="list-style-type: none"> • Strategies • Tactics • Rule Sets • Orders • TPFDD | Control <ul style="list-style-type: none"> • MOEs • Model • Control |
|--|---|---|---|---|



UNCLASSIFIED

Joint Data
Support

UNCLASSIFIED

What's an Analytical Baseline?

Definition (DODD 8260.1): A package comprising a scenario, concept of operations, and integrated data used by DoD Components as a foundation for strategic analyses.

- | | | |
|--|---|--|
| Scenario <ul style="list-style-type: none"> • Location • Year • Objectives | CONOPS <ul style="list-style-type: none"> • Red • Blue | Other Data <ul style="list-style-type: none"> • Characteristics and Performance (C&P) • Forces, Units, and Equipment (FUE) • Environmental (terrain, weather, ...) |
|--|---|--|

Plus Analysis

To encourage internal consistency

If models are used in the analysis it will include model input data, metadata, the model (if available), and output results

Goal: Visible, accessible and traceable data ready to support Department-level analyses

UNCLASSIFIED

UNCLASSIFIED



Definitions

- **Analytical Baseline.** A package comprising a scenario, concept of operations, and integrated data used by the DoD Components as a foundation for strategic analyses. Examples of analytical baselines include scenarios and supporting data used for computer-assisted war games and theater campaign simulations.
- **Concept of Operations.** A verbal or graphic statement, in broad outline, of a commander's assumptions or intent in regard to an operation or series of operations. The concept of operations frequently is embodied in campaign plans and operation plans, in the latter case particularly when the plans cover a series of connected operations to be carried out simultaneously or in succession. The concept is designed to give an overall picture of the operation. It is included primarily to clarify the purpose of the operation. Also called commander's concept or CONOPS.
- **Current Forces.** Forces in existence in the current fiscal year (the execution year) or planned to be available in the first year of the Future Year Defense Program (FYDP).

UNCLASSIFIED

UNCLASSIFIED



Definitions (continued)

- **Data.** Representations of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by human or automatic means. Any representations, such as characters or analog quantities, to which meaning is or might be assigned.
- **Data Repository.** An organized collection of data and metadata, suitable for use in developing analytical baselines, strategic analyses, or associated products.
- **Data Verification and Validation.** The process of substantiating the internal consistency and accuracy of data and certifying that the data represent real-world entities and are appropriate for their intended purpose or expected range of purposes. When carried out by data providers, the term "data validation" refers to the documented assessment of data by subject-area experts and the comparison of the data to known values.
- **Force.** An aggregation of military personnel, weapon systems, equipment, and necessary support, or a combination thereof.

UNCLASSIFIED



UNCLASSIFIED

Definitions (continued)

- **Future Forces.** Forces projected to be available beyond the first year of the FYDP.
- **Joint Analytical Baseline Development Team.** An ad hoc team of representatives from selected DoD Components, charged with creating an analytical baseline.
- **Metadata.** Information describing the characteristics of data; data or information about data; and descriptive information about an organization's data, data activities, systems, and holdings.
- **Scenario.** An account or synopsis of a projected course of action or events. For purposes of this Instruction, the focus of scenarios is on strategic and operational levels of warfare. Scenarios include information such as politico-military contexts and/or backgrounds, assumptions, operational objectives (threat and friendly), major force arrivals, and planning considerations.

UNCLASSIFIED



UNCLASSIFIED

Definitions (continued)

- **Strategic Analysis.** An analysis of force sufficiency and effectiveness conducted by the DoD Components to support the development and evaluation of the defense strategy. Such analyses address both forces and enablers (e.g., inter-theater and intra-theater lift capability).
- **Strategic Analysis of Current Forces.** An analysis of the sufficiency and effectiveness of current forces, including enablers. Such analyses examine the current force's ability to execute the defense strategy. These analyses inform longer-term strategic analyses.
- **Strategic Analysis of Future Forces.** An analysis of the sufficiency and effectiveness of out-year forces, including enablers. Such analyses examine force structure and program alternatives and evaluate risks in the midterm and beyond.

UNCLASSIFIED

UNCLASSIFIED



Joint Data Support
Crystal Gateway 2, Suite 200
1225 Jefferson Davis Highway
Arlington, VA 22202
UNCLAS fax: (703) 414-8114
SIPRNet: jds@pac.osd.smil.mil
SIPRNet website: <http://jds.pac.osd.smil.mil>

-
- Dr. Jim Stevens, Director, (703) 601-0406, james.stevens@osd.mil
 - Eric Johnson, Study Support, (703) 601-0411, richard.e.johnson@osd.mil
 - Rick Jimenez (U.S.) / Dan Lotz (non-U.S.), Data Acquisition & Analysis, (703) 604-6384 (both), ricardo.jimenez@osd.mil, dan.lotz@osd.mil
 - JP Wilusz, IT Support, (703) 414-1959, jscph.paul.wilusz@osd.mil
 - Zeke Dennison, JDS JWARS Team, (703) 696-9490, ezekiel.dennison@osd.mil
 - Randy Wimmer, JDS Scenario Data Support Team, (703) 414-1966, james.wimmer@osd.mil

UNCLASSIFIED

Acronyms

Improving Defense Analysis
Through Better Data Practices
March 25 - 27, 2003, Alexandria, VA

ABCA	American, British, Canadian and Australian
ADS	Authoritative Data Sources
AF/ASC	Air Force Aeronautical Systems Center
AFIT	Air Force Institute of Technology
AFMIC	Armed Forces Medical Intelligence Center
AFSAA	Air Force Studies and Analysis Agency
AFSMC	Air Force Space and Missile Command
AMSAA	Army Materiel System Analysis Agency
AoA	Analysis of Alternatives
ARI	Army Research Institute
ASC	Aeronautical Systems Center
ASD(C3I)	Assistant Secretary of Defense (Command, Control, Communications and Intelligence)
BRAC	Base Realignment and Closing
C2	Command and Control
C3I	Command, Control, Communications and Intelligence
C4ISR	Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance
CAA	Center for Army Analysis
CAC	Common Access Card
CASCOM	Combined Arms Support Command
CC	Command Center
CHEMTREC	Chemical Transportation Emergency Center
COBP	Code of Best Practices
CoI	Community of Interest
CONOPS	Concept of Operations
COTS	Commercial Off The Shelf
CP&A	Capability Planning and Analysis
CWM	Common Warehouse Metamodel
DCI	Director of Central Intelligence
DDMS	DoD Discovery Metadata Standard
DHS	Department of Homeland Security
DIA	Defense Intelligence Agency
DISA(C ³)	Defense Information Systems Agency
DLA	Defense Logistics Agency
DM	Data Management
DMSO	Defense Modeling and Simulation Office
DoD	Department of Defense
DODD	Department of Defense Directive
DODI	Department of Defense Instruction
E&T	Education and Training

EBO	Effects Based Operations
FFRDC	Federally Funded Research and Development Center
FS	MORS Fellow of the Society
GOTS	Government Off The Shelf
HIU	Humanitarian Information Unit
HLA	High Level Architecture
HLS	Homeland Security
HQ	Headquarters
HQDA	Headquarters Department of the Army
HQDA/G3	Headquarters Department of the Army, Operations
IO	Information Operations
J-4	The Joint Chiefs of Staff, Logistics
J-8	The Joint Chiefs of Staff, Force Structure
JADM IPT	Joint Analytic Data Management Integrated Process Team
JADMSC	Joint Analytic Data Management Steering Committee
JDS	Joint Data Support
JFCOM	US Joint Forces Command
JLA	Joint Logistics Agency
JMEM	Joint Munitions Effectiveness Manual
JSCAP	Joint Strategic Capabilities Plan
JSF	Joint Strike Fighter
JTIMS	Joint Training Information Management System
JWARS	Joint Warfare System (DoD's newest theater level model)
M&S	Modeling and Simulation
MCCDC	Marine Corps Combat Development Command
MCIA	Marine Corps Intelligence Activity
METL	Mission Essential Task List
METL	Mission-Essential Task List
MOF	Meta Object Facility
MOP	Measures of Performance
MOUT	Military Operations in Urban Terrain
MSIC	Missile and Space Intelligence Center
MSRR	Modeling and Simulation Resource Repository
N-81	Assessment Branch of Deputy Chief of Naval Operations (Resources, Requirements and Assessments)
NAIC	National Air Intelligence Center
NATO	North Atlantic Treaty Organization
NCES	Net Centric Enterprise Services
NDIA	National Defense Industrial Association
NDU	National Defense University
NGIC	National Ground Intelligence Center
NGO	Non-Governmental Organizations
NIH	National Institutes of Health
NOAA	National Oceanographic and Atmospheric Agency
NPS	Naval Postgraduate School
NSF	National Science Foundation

NSTC	National Science and Technology Council
NTC	National Training Center
OASD(C3I)	Office of the Assistant Secretary of Defense (C3I)
ODS	Output Data Set or Optical Data Storage
OEF	Operation Enduring Freedom
OLAP	On-Line Analytical Process
ONI	Office of Naval Intelligence
OOTW	Operations Other Than War
OSD	Office of the Secretary of Defense
OT&E	Operational Test and Evaluation
OWL	Ontology Web Language
PA&E	Program Analysis and Evaluation
PKI	Public Key Infrastructure
PMs/RDECs	Program Managers and (the Army's) Research, Development and Engineering Center
POA&M	Plan of Action and Milestones
PPBS	Planning, Programming and Budgeting System
QDR	Quadrennial Defense Review
R&D	Research and Development
RDF	Raw Data File
RTO	Research and Technology Organization (NATO)
SAML	Security Assertion Markup Language
SAR	Special Access Required
SAS	NATO Studies, Analysis and Simulations
SIMTECH	Simulation Technology, a MORS Workshop in 1997
TMAP	Threat Model Assessment Process
TOA	Table of Allowance
TPED	Task, Process, Exploit, Disseminate
TPPU	Task, Post, Process, Use
TRAC	TRADOC (Training and Doctrine Command) Analysis Center
TRANSCOM	US Transportation Command
TTCP	The Technical Cooperation Program (Treaty arrangement by the United States, United Kingdom, Canada, Australia and New Zealand that permits the exchange of classified technical information under certain conditions)
UN	United Nations
UOB	Unit Order of Battle
USAF	United States Air Force
USMC	United States Marine Corps
USN	United States Navy
V&V	Verification and Validation
VISTA	Visualized Information and Synthesized Temporal Analysis
VV&A	Verification, Validation and Accreditation
WMD	Weapons of Mass Destruction
XMI	XML Metadata Interchange
XML	Extensible Markup Language
XSD	XML Schema
XSLT	XSL Transformations

Improving Defense Analysis Through Better Data Practices

A Workshop

Terms of Reference

1. Background

Effective quantitative analysis depends on the availability of credible data. Each of the Services, the Joint Staff, and the many organizations within the Office of the Secretary of Defense have established a variety of procedures to gather, generate, maintain, transform and use military data to support an array of analytic activity, from budget programming, to strategic planning, to wargaming. In fact, the collection and application of credible data is a major part of every analytic effort and a key contributor to the value of every final product. In order to better support strategic analyses conducted by the Department, *Defense Planning Guidance, Fiscal Years 2004-2009, May 2002*, tasked the Director of Program Analysis and Evaluation to establish policy and assign responsibility for generating, collecting, developing, maintaining, and disseminating data associated with current and future U.S. and non-U.S. forces. Existing directives and pending instructions provide general policy for accomplishing this task. To manage, gather, create and apply analytic data, the Services and defense organizations currently employ a wide variety of methodologies, many resulting from several years of evolution. At the same time, commercial enterprises and the academic community have generated many new alternative methodologies for handling data within their disciplines to include new technologies that could be applicable to the military. This paper describes a Military Operations Research Society special meeting to review current data processes used by DoD to support analysis, identify best data practices throughout the DoD and broader community, and recommend, where appropriate, best practices that could be implemented to improve the analytic data processes used by the Department.

2. Discussion

Recent MORS special meetings have addressed a myriad of analytic focus areas such as weapons of mass destruction and urban operations; new analytic tools, to include the new sciences; and, in the near future, the shaping of analytic personnel to support the needs of the combatant commanders. Key to each of these areas and noted in most special meeting final reports is the importance of good data to the usefulness of the supporting analysis and the credibility of the final products. Directives within the Department provide general guidelines for data verification and validation (data V&V) to support the range of analyses conducted. These directives do not include guidance on specific methodologies, techniques, tools, procedures or the use of metadata that would help ensure data is of the appropriate quality and used in the appropriate ways to best contribute to the credibility of the final product. In fact, some study results are undermined by the lack of credible data to support the analysis, while others are marred by the misapplication of credible data to models, scenarios, and methodologies that were outside the scope the data was designed to support. In fact, concerns regarding the collection and application of warfighting databases have often served to impede the

production of credible joint warfighting analyses. For example, program offices carefully develop system lethality and vulnerability results to address specific threats in predetermined contexts in order to produce estimates of military utility associated with the fielded system. These data may or may not be appropriate for use in joint warfighting analyses addressing the effectiveness of an array of systems operating jointly against an array of threats that were outside the scope of the original data development process. While the problems may not be due to any intent by the analysts involved to skew the estimated effectiveness results, misapplications and misrepresentations still occur. As a result, concerns about the use of the data have led to barriers within the Department to developing balanced, objective and credible databases and analyses.

These issues are not unique to DoD or to the various Service elements. Non-DoD government agencies (such as the Bureau of Labor Standards), commercial business organizations, other governmental and inter-governmental organizations (to include the World Bank) and the academic community in general deal with similar issues across all elements of enterprise activity. These groups, as well as many agencies within the military, have developed specific methodologies to assist in the production of credible databases in order to limit unintended biases and misapplication of data within general warfare analysis areas or across other applications. MORS is in a unique position to help address these data production, management, and application issues by inviting analysts, policy makers, database managers and experts in the field to review existing procedures and then to identify and recommend best data practices for military and joint warfighting analysis to the Department. These best practices can then be used as a standard to guide more specific policies within the Department of Defense, particularly with respect to supporting strategic analyses, as well as to help individual units gather and maintain credible data for specific purposes. Such a special meeting will support current plans for the Under Secretary of Defense for Policy to coordinate the development and/or updating of DoD scenarios, as well as for other activities within the Department to ensure the availability of credible data for a range of military activities.

3. Goals and Objectives

The purpose of this workshop would be to identify best practices for generating, collecting, developing, maintaining, disseminating and applying data and metadata to help support better analyses as well as to help establish improved data policy and procedures within the Department of Defense. In particular, the workshop will examine current practices by organizations both inside the Department and throughout the inter-agency, international, commercial and academic worlds to determine those that incorporate best data practices. In addition, the workshop would assess, to the extent practical, new technologies to assist in the management and application of data. As a part of this activity, the workshop would address potential tools, techniques, procedures and technologies for identifying, collecting, verifying, managing, and disseminating data and analytical baselines within the Department. Specific best practices will be identified regarding:

- What data content would best support the development of analytic baselines (defined by DODD 8260.1 as a package comprising a scenario, concept of operations, and integrated data used by DoD Components as a base case for

strategic analyses) to include methods of collecting and storing data, metadata, assumptions, scenarios contexts and other pertinent information?

- What cost effective methods and approaches best support verification and validation checks on data received, to include associated metadata in data repositories?
- What techniques are available to ensure data is appropriate for its intended uses be they extensions of existing applications or new applications?
- What effective and efficient methods are available for preparing data sets and associated metadata for use in specific studies and analyses to conclude the configuration of related data sets, delivering data according to criteria established by oversight bodies and archiving data and other study products for later reporting and application?
- What guidelines and methodologies are available for prioritizing data development, storage, management, and dissemination activities?

In addition, to the extent possible, the workshop will identify organizations and activities within the Department that are perceived to be authoritative data sources as well as identify characteristics that help define an authoritative data source, including the internal practices processes and procedures that help ensure data credibility.

3.1 Goals.

This workshop will afford the military OR community an opportunity to achieve the following goals:

- (1) Establish a vision for improving the generation and management of credible data to support the analytic activities of the Department;
- (2) Identify specific best practices that could be adopted by the Department, the Joint Staff, the Services, and their supporting agencies to realize this vision;
- (3) Define how MORS can work with the sponsors over time to implement improvements to data production and management.

3.2 Objectives.

The objectives of the workshop are to:

- Review current DoD processes for developing, managing, using, and sharing data in support of military analyses to include defining common terms
- Identify best practices and standards within DoD and the broader analytic community for generating, collecting, developing, maintaining, and disseminating data; suggest how applicable best practices could be implemented in DoD
 - Review expected impact of technology on practices and processes
 - Note issues and problems that serve as barriers to implementing best practices

- Identify and assess potential sources of data needed for a wide variety of analyses
 - Provide a current listing of authoritative data sources within the Department
- Educate the community on the new Data Directive, to include plans for implementation (data repository, access to the data, etc.)

4. Approach

In order to achieve these goals and objectives, the subjects of data management, data generation, data creation, and data transformation and application will be examined thoroughly. General presentations by Department and Service experts regarding current practices and desired best practices will be followed by descriptions of methodologies that have been developed and implemented by other government agencies, international and/or allied organizations, industry, and academia. Working groups will address data practices within specific activity and contextual frameworks. Working groups will be asked to review existing practices in their area within DoD as a starting point and then will discuss relevant practices by non-DoD organizations as well as techniques and approaches addressed in the literature. By developing a vision of the desired outcome in each area, the working groups will then identify and recommend efficient and effective practices in their areas for consideration by the Department. The working group focus areas, selected issues, and desired results are detailed below.

4.1 Working Groups.

Following general presentations, the workshop attendees will be organized into five subject area working groups plus a synthesis working group. Depending on the numbers of the participants, additional groups can be generated to address issues in more depth with different individuals in order to provide a broader range of best practice options for the Department. The focus will be on identifying best data practices as well as barriers and procedures that inhibit the ability to manage, gather, create, share, transform and apply credible data as well as on technologies developed to support these activities. Specific working groups will be as follows:

- Data Management
- Obtaining Data From Outside Sources
- Creating Data Internally
- Transforming and Employing Data
- Data Support Technology
- Synthesis

4.1.1 Data Management Working Group.

This working group will review current DoD data management practices and investigate data management approaches by other pertinent government organizations, such as the Bureau of Labor Statistics, as well as industry and international organizations to determine their applicability to managing data for military analysis. The group will

identify best community practices of data management, to include enterprise business rules, models, and structures to facilitate effective data management and data sharing. In addition, the group will review storage; access; pedigree; validation and verification (V&V) requirements; currency; tools; preservation of data context, and the use of meta data. The group will suggest ways to implement best practices if appropriate. In addition, the group will address the handling of critical technical and compartmentalized data, and suggest generic ways for addressing classification issues in support of joint analytic processes. The group will produce a summary briefing report identifying specific best practices for data management by the Department and its organizations and agencies.

4.1.2 Obtaining Data from Outside Sources Working Group.

This "Data Demand" group will review current data gathering and data sharing processes within DoD as well as "best practice" data gathering and data sharing approaches in the larger community. Such practices could include composable frameworks and supporting procedures; tools, utilities and supporting standards; critical mass of end-use content and supporting access; and life-cycle sustainment and supporting elements. The group will address red and blue data issues such as release, responsibilities and liabilities and suggest ways to improve the process. To the extent possible, the group will identify currently perceived authoritative sources for general data types. The group will also identify databases that are vital to support the current OSD Analytic Agenda, but that are either of poor quality or not available at all. To the extent possible, the group will provide possible solutions that could remedy this situation in a timely fashion. The group will produce a summary briefing report identifying best data gathering practices, along with current impediments to effective data collection and recommended approaches to facilitate the adoption of best practices by Department organizations and agencies.

4.1.3 Creating Data Internally Working Group.

This "Data Provider" working group will review current Department and community practices for data creation, to include the expected attributes of an authoritative data source. It will identify the data content associated with various types of data produced to support joint studies, experiments and wargames, list potential sources for this data, understand and evaluate how data is developed by various sources, and identify best practices in this arena. The group will also address the implications of the Intelligence community's transition to Matlab/Simulink representations for threats. Finally, the group will suggest ways to incorporate best practices in the data creation process, as appropriate. Particular attention will be paid to specific techniques that could facilitate sharing of effective approaches to data creation across defense organizations. These will include cost- and analytic resource-effective methods and procedures to ensure the appropriate application of one Service's or Organization's or Program Office's data to issues and studies conducted by other defense organization. The group will produce a summary briefing report identifying data content, potential sources, evaluation of data creation, and the identified best practices, along with current impediments to credible data creation and recommended approaches to facilitate the adoption of best practices by Department organizations and agencies.

4.1.4 Transforming and Employing Data Working Group.

This "Data Application" working group will review current DoD and best practices of data application, analysis and sharing to include appropriate use of data and maintaining data integrity when aggregating, disaggregating and transforming data for different applications. The group will suggest ways to implement best practices if appropriate. In its investigation, the group will review best practices from other government agencies, pertinent international or allied organizations, industry and DoD business applications. The examination will include discussion of how data is currently applied to department strategic analyses and what methods are available to assure the best application of that data to both traditional and new analytic problem domains. In particular, this group will identify specific tools and procedures to help ensure the appropriate application of existing data or the appropriate transformation of data from existing databases to best support issues confronting the joint community. The group will produce a summary briefing report describing the data transformation and application methods and approaches explored, the best practice applications that may be addressed with known techniques, and those applications that would be better served by the development of new techniques. The briefing report will also list current impediments to credible data application and suggest recommended approaches to facilitate the adoption of best practices within the Department.

4.1.5 Data Technology Support Working Group.

This working group will address state of the art technology to support the needs of the other data enterprise areas of management, collection, creation, transformation and employment. The group will identify and assess the impact and appropriate use of new technologies in areas such as data composition and data mining. The group should pay particular attention to methodologies used to track databases and database changes to ensure traceability throughout the analytic process as well as cost-effective methods to ensure reasonable levels of data verification and validation are maintained. Discussions will also include data engineering and data interchange standards. In addition to addressing current DoD practices, the group will identify best practices used by industry, other government agencies, and international organizations. The group will produce a summary report identifying best practice data support technologies, along with current impediments to implementation of these technologies and recommended approaches to facilitate their adoption by Department organizations and agencies

4.1.6 Synthesis Working Group

This working group will take a broad view, identifying the high level issues across the domains of data generation, management, and application that enable or impede the successful evolution of joint and warfighting and strategic analysis. The group will provide the integrating perspective to help highlight best data practices, particularly in those areas where techniques that reach across all areas will best meet the long-range requirements of the Department. This working group will provide the integrating perspective.

4.2 Working Group Tasking.

Working groups will be directed to address the following questions and concerns:

How do the Department and its Services and organizations currently perform the functions of the specific area that it is addressing (data management, collection, creation, or application)?

How are these functions performed by other enterprises in other government departments, commercial enterprises, international organizations or by the scientific and academic communities?

What features of existing or proposed alternative methods, approaches and techniques appear to be best suited for adoption and implementation within the Department?

Working groups will identify practices that help or hinder operations within their focus area and then will define the key best practices that they feel would have the greatest impact on improving data management, data gathering, data creation, data transformation, data application and sharing within the Department. Each "best practice" will be defined in sufficient detail to help potential DoD users to understand how the method could be implemented and what metrics could be used to measure progress toward full implementation at both the organizational and department level. Time permitting, a proposed plan for implementing the most important practices will also be developed along with the likely impact of current or emerging technologies on best data practices in the previously described areas.

5. Sequence of Events.

A warm-up session for working group co-chairs will be held the evening before the first day of the workshop. This session will be held at the workshop hotel. The purpose of this warm-up session is to review and discuss tasking for the working groups.

The first day will be devoted to a plenary session. The Workshop Co-Chairs will present the goals, objectives, and organization for the workshop. A MORS Sponsor will present a detailed charge to the workshop participants. Session speakers will provide alternative perspectives on the current state of data generation, management, application and sharing for joint warfighting analysis within the Department; as well as best data practices from other government and commercial enterprises and from the scientific/academic world.

There will be a social event (mixer) at the end of the first day.

All of the second day will be devoted to working group sessions and discussions. The morning of the third day will be devoted to the preparation of working group presentations. The afternoon of the third day will be devoted to working group presentations in plenary session.

The morning of the fourth day will be made available to working group co-chairs for final editing of the working group annotated briefings.

6. Agenda

Day/Time	Activity	POC	Location
Monday 1700	March 24, 2003 Working Group Co-Chair Warm-Up Session	Workshop Co-Chairs	

Tuesday	March 25, 2003	
0700	Registration and Continental Breakfast	MORS Office
0800	MORS President's Welcome	Ted Smyth
0805	Facility Host Welcome	Phil Major
0810	Sponsor's Welcome	Eric Coulter
0815	Chair's Welcome, Workshop Overview	Tom Allen
0830	Keynote, DoD Data Enterprise	John Osterholz
0910	Keynote, DoD Analytic Agenda	Chris Lamb
0950	Break	
1000	OSD Enterprise	Tony Simon
1030	DoD Data Directive & Analytic Agenda Implementation	Jim Stevens
1100	Service/JS Perspectives	Senior Analyst Panel Discussion
1230	Buffet Lunch	
1330	Industry Standards for Data Generation	Invited Speaker, Industry
1415	Non-DoD Data Management	Kurt Schnebele, NOAA
1500	Break	
1515	Best Practices in Data Applications/Sharing	Invited Speaker
1600	Best Practices in Data	Invited Speaker
1645	Working Group Kick-Off Meetings	Working Group Chairs
1730	Mixer	MORS Office
Wednesday	March 26, 2003	
0715	Continental Breakfast	MORS Office
0800	Working Group Session #1	Working Group Co-Chairs
0945	Break	
1000	Working Group Session #2	Working Group Co-Chairs
1130-1300	Staggered Lunch, IDA Cafeteria:	MORS Office
1300	Working Group Session #3	Working Group Co-Chairs
1500	Break	MORS Office
1515	Working Group Session #4	Working Group Co-Chairs
1700	Working Group Chair & Co-Chairs Hotwash	Workshop Chair
Thursday	March 27, 2003	
0715	Continental Breakfast	MORS Office
0800	Working Groups Session #5	Working Group Co-Chairs
0945	Break	MORS Office
1000	Working Groups Session #6 (Prepare Briefing)	Working Group Co-Chairs
1200-1330	Staggered Lunch, IDA Cafeteria	MORS Office
1330	Working Groups: Present Briefings, WG 1, 2, 3	Working Group Co-Chairs
1500	Break	MORS Office
1515	Working Groups: Present Briefings, WG 4, 5 and Synthesis Group	Working Group Co-Chairs
1645	Workshop Wrap-Up	Workshop Chair
1700	Adjourn	Workshop Chair
Friday	March 28, 2003	
0800	Working Group Co-Chairs complete Working Group Annotated Briefing	Working Group Co-Chairs
1200	Adjourn Workshop	Workshop Chair

7. Attendees

Attendance will be controlled via invitation. Attendees will include invited experts from OSD, all Services, the Joint Staff, Federally Funded Research and Development Centers, operational commanders, DoD contractors, and data base developers and managers from other government departments, commercial firms and academia. Workshop chairs will control membership of their sessions in conjunction with the Organizing Committee. Attendance will be limited to 150 people.

8. Products

There will be six specific products generated as a result of this workshop:

- A collection of common terms and definitions associated with data management, application and generation available for use by the military analysis community. An attempt will be made to generate a draft of this document prior to the workshop, but it will be updated and finalized as a part of the workshop process.
- An integrated list of data best practices in the areas addressed.
- An Executive Summary of the workshop in the form of a text document and a scripted briefing for the MORS Sponsors addressing the workshop objectives, findings, conclusions, and recommendations.
- A proceedings document containing summaries of all sessions and copies of appropriate briefing slides and presentations.
- A general session presentation for the 71st MORSS.
- A *PHALANX* Article.

9. Planning and Organizing Chairs and Committees

General Chair	Tom Allen
Co-Chair	Jim Bexfield
Plenary Facilitator	Scott Simpkins
Data Management WG	Simone Youngblood, Jim Stevens
Data Gathering WG	Jack Sheehan, Phillipe Loustaunau
Data Creation WG	Sam Fragapane, Bryan Paris, Dennis Leedom
Data Employment WG	Clay Bowen, Ernie Boehner, Rudy Pabon
Data Technology WG	Bob Might, Ron Smits Jim Richardson
Synthesis WG	Stu Starr, Roy Reiss, Kirk Yost
Admin Coordination	Brian Engler, Natalie Kelly
MORS Bulldog	Bill Dunn

OSD Rep	Jim Bexfield
Joint Staff Rep	Bob Orlov
Army Rep	Paul Deitz
Navy Rep	Herb Cupo
Air Force Rep	Roy Reiss
Marine Corps Rep	Phil Exner
DMSO	Simone Youngblood
MORS	Lee Dick
Other	

10. Administration

Name: Ms. Natalie Kelly, MORS, 1703 N. Beauregard St, Suite 450, Alexandria, VA 22311

Dates: 25-27 March 2003

Location: IDA facility, Alexandria VA

Fee: \$210 government, \$420 all others

Attendance: 120 - 150